

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: September 25, 2004, 01:21:23 ; Search time 1098 Seconds

(without alignments)
8264.254 Million cell updates/sec

Title: US-10-010-720-14

Perfect score: 10812

Sequence: 1 MSGGAEXKXSTPGSLFLSP.....NISNLQKISINPPGSNLTFTT 2136

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:
-MODEL=frame_p2n.model -DEV=xlh
-Q=/cgn2_1/USPRO.spool/US10010720/runat_23092004_154319_25232/app.query.fasta_1.2311
-DB=N_Geneseq_29Jan04 -OPT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPCL=0
-LOOEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blonum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=PCT -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NOEM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USPR=US10010720 @CGN 1.1 818 -runat_23092004_154319_25232 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERRY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseq_29Jan04:*

1: geneseqn1980s:*\n2: geneseqn1990s:*\n3: geneseqn2000s:*\n4: geneseqn2001as:*\n5: geneseqn2001bs:*\n6: geneseqn2002s:*\n7: geneseqn2003as:*\n8: geneseqn2003bs:*\n9: geneseqn2003cs:*\n10: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	10810	100.0	6411	8	ACH03770 DNA encod
2	10794.5	99.8	7280	6	AAD34306 Human PKI
3	10753.5	99.5	6690	8	ACH03767 DNA encod
4	10677	98.8	7149	8	ABZ59199 Human WNK
5	10677	98.8	7149	8	ACH03764 DNA encod
6	10646	98.5	6327	8	ACH03778 DNA encod
7	10646	98.5	7152	9	ADG99101 Human KRP
8	10589.5	97.9	6606	8	ACH03775 DNA encod

9	10513	97.2	7065	8	ACH03772 DNA encod
10	10508	97.2	6231	8	ACH03786 DNA encod
11	10451.5	96.7	6510	8	ACH03783 DNA encod
12	10375	96.0	6969	8	ACH03780 DNA encod
13	10344	95.7	6147	8	ACH03794 DNA encod
14	10287.5	95.1	6426	8	ACH03791 DNA encod
15	10211	94.4	6885	8	ACH03788 DNA encod
16	9929	91.8	6000	8	ACH03771 DNA encod
17	9872.5	91.3	6279	8	ACH03768 DNA encod
18	9796	90.6	6738	8	ACH03765 DNA encod
19	9765	90.3	7596	8	ACH03779 DNA encod
20	9729.5	90.0	7328	4	AA44683 Novel pro
21	9708.5	89.8	6195	8	ACH03776 DNA encod
22	9632	89.1	6654	8	ACH03773 DNA encod
23	9627	89.0	5820	8	ACH03787 DNA encod
24	9570.5	88.5	6099	8	ACH03784 DNA encod
25	9494	87.8	6558	8	ACH03781 DNA encod
26	9463	87.5	5736	8	ACH03795 DNA encod
27	9406.5	87.0	6015	8	ACH03792 DNA encod
28	9330	86.3	6474	8	ACH03789 DNA encod
29	4053	37.5	2949	8	ACH03766 DNA encod
30	4043	37.4	2490	8	ACH03769 DNA encod
31	3889	35.0	2865	8	ACH03774 DNA encod
32	3879	35.9	2406	8	ACH03782 DNA encod
33	3751	34.7	2769	8	ACH03785 DNA encod
34	3741	34.6	2310	8	ACH03780 DNA encod
35	3587	33.2	2685	8	ACH03793 DNA encod
36	3577	33.1	2226	8	ACH03793 DNA encod
37	3552.5	32.9	2396	4	AA157892 Human pol
38	3498	32.4	2663	4	AA157893 Human pol
39	2953	27.3	2310	5	AAH76212 Human kin
40	2859.5	26.4	19167	4	AA105340 Human rep
41	2859.5	26.4	19167	4	AB198209 Human tes
42	2859.5	26.4	19167	5	ABAI9101 Human ner
43	2642	24.4	7825	7	ACC57741 Human pro
44	2627	24.3	1536	4	AA159679 Human pro
45	2627	24.3	1536	4	AA159678 Human pol

ALIGNMENTS

RESULT 1
ACH03770
ID ACH03770 standard; DNA; 6411 BP.

XX AC ACH03770;

XX DT 26-SEP-2003 (first entry)

XX DE DNA encoding novel human protein kinase #7.

XX KW Human; kinase; ds; gene; cosmetic application; nutraceutical application.

XX OS Homo sapiens.

XX PN US6541252-B1.

XX PD 01-APR-2003.

XX PF 14-MAY-2001; 2001US-00854856.

XX PR 19-MAY-2000; 2000US-0206015P.

XX PA (LEXI-) LEXICON GENETICS INC.

XX PI Walke DW, Hilbun E, Donoho G, Turner CA;

XX DR WPI, 2003-575927/54.

XX DR P-PSDB; ABO44387.

XX PT New nucleic acid encoding novel human proteins, useful in cosmetic and nutraceutical applications.

QY 621 AlaSerValSerThrGlnValGluProGluGluProGlnAlaAspGlnHisGlnGlnLeu 640
DB 1861 GCTTCAGTTCTTCAACAGTAGAAGCTGAAGAACCTGAGGACAGATCAACATCAACAACCTA 1920
QY 641 GlnTyrGlnGlnProSerTlleSerValLeuSerAspGlyThrValAspSerGlyGlnGly 660
DB 1921 CAGTACCAAGCAACCGATATATCTGTGTATCTGATGGAGGCTGACATGTGTACAGGGA 1980
QY 661 SerSerValPheThrGlnSerArgValSerSerGlnGlnThrValSerTyrGlySerGln 680
DB 1981 TCTCTGTCTTCCAGCAATCTCAGTAGACAGCCACACAGACATTTCAATGTGTTCCCA 2040
QY 681 HisGlnGlnAlaHisSerThrGlyThrValProGlyHisTlleProSerThrValGlnAla 700
DB 2041 CATTGAACAGGCACATTTTCAAGGACAGCTCCAGGGCATATACCTTCTACTGTCCAGCA 2100
QY 701 GlnSerGlnProHisGlyValTyrProProSerSerValAlaGlnGlyGlnSerGlnGly 720
DB 2101 CAGTCTCAGCCCATGGGTATATCCACCTCAAGTGTGCACAGGGGCGAGCCAGGGT 2160
QY 721 GlnProSerSerSerSerLeuThrGlyValSerSerSerGlnProTlleGlnHisProGln 740
DB 2161 CAGCATCTTCAAGTAGCTTAAACAGGGGTTTCACTTCCACACCATTAACATCTCTCAG 2220
QY 741 GlnGlnGlnGlyTlleGlnGlnThrAlaProProGlnGlnThrValGlnTyrSerLeuSer 760
DB 2221 CAGCAGAGGGAATACAGCAGACAGCCCTCTCAACAGACAGTACAGTATTCACCTTCA 2280
QY 761 GlnThrSerThrSerSerGlyAlaThrThrAlaGlnProValSerGlnProGlnAlaPro 780
DB 2281 CAGACATCAACCTCCAGTAGAGCCACATACGACAGCAGTAGATCCAGCTCAAGCTCCA 2340
QY 781 GlnValLeuProGlnValSerAlaGlyValGlnSerThrGlnGlyValSerGlnValAla 800
DB 2341 CAAATCTTGGCTCAAGTACGCTGGAAGAACAGAGTCTCAGGGAGTCTCTCAGGTTGCT 2400
QY 801 ProAlaGluProValAlaValAlaGlnProGlnAlaThrGlnProThrThrLeuAlaSer 820
DB 2401 CTTCCAGACGCAATGGCAGTAGACACAGCCCAAGCTTCCAGCGGACACATTTGGCTTCC 2460
QY 821 SerValAspSerAlaHisSerAspValAlaSerGlyMetSerAspGlyAsnGlnAsnVal 840
DB 2461 TCTGTACAGAGTGCATTCAGATGTTGCTTCAAGTATGATGATGCAAGAACAGTCTC 2520
QY 841 ProSerSerSerGlyAspGlnGlyValArgThrThrIleAspArgHisTyrArgGlySerVal 860
DB 2521 CCATCTTCCAGTAGAAGGACATGAAGAGAACCTACAAACCGCATACCGAAATCTGTA 2580
QY 861 ArgSerArgSerArgHisGlnGlyValSerArgProIleuAspArgIleLeuAsnValSer 880
DB 2581 AGGAGTCCCTCTCGACATGAABAAAATTTCAAGCCCAAAATTTAGAAATTTGAAATGTTCA 2640
QY 881 AsnIleGlyAspArgValValGluGlySerGlnLeuGluThrHisAsnArgGlyMetValThr 900
DB 2641 AATAAAGGAGACCCAGTAGTGAATGTCAATTAGAGACTCAATAATAGAGAAATGTTACA 2700
QY 901 PheIlePheAspLeuAspGlyAspAsnProGluGluIleAlaThrIleMetValAspAsn 920
DB 2701 TTCAAATTTGACCTTAGATGTGACAAACCCAGAGGATGACAAACATTTATGTGACAAAT 2760
QY 921 AspPheIleLeuAlaIleGluArgGluSerPheValAspGlnValArgGluIleIleGlu 940
DB 2761 GACTTTATTTCTAGCAATAGAGAGAGCTTTTGTGATTCAGAGTCGAGAAATTTATTTAA 2820
QY 941 LysAlaAspGluMetLeuSerGlyAspValSerValGluProGluGlyAspGlnGlyLeu 960
DB 2821 AAAGCTATGAATCTCAGTAGAGATGTCAGTGTGAAACAGAGGGGTGATCAGGGATTTG 2880
QY 961 GlnSerLeuGlnGlyValAspAspTyrGlyPheSerGlySerGlnIleLeuGlnGlyGln 980
DB 2881 GAGAGTCTTACAGAGAAAGATACATATGCTTTTCAAGTCTTCAAAATTTGAGAGAGAG 2940
QY 981 PheIleGlnProIleProAlaSerSerMetProGlnGlnIleGlyIleProThrSerSer 1000

DB 2941 TTCAAACCAACCAATTCCTGCTTCCATGCCACACAAATAGCATTCCTTACCAATCTCT 3000
QY 1001 LeuThrGlnValAlaHisSerAlaGlyArgAspPheIleValSerProValProGluSer 1020
DB 3001 TTTACTCAAGTTCTTCAATCTCTGCGGAGAGCGGTTTATATGATGCTCTGTGCCAGAAAGC 3060
QY 1021 ArgLeuArgGluSerLysValPheProSerGluIleThrAspThrValAlaAlaSerThr 1040
DB 3061 CGATTACGAGAAATCAAAAGTTTCCCAAGTAAATTAACAGATACAGTTGCTGCTCTTACA 3120
QY 1041 AlaGlnSerProGlyMetAsnLeuSerHisSerAlaSerSerLeuSerLeuGlnGlnAla 1060
DB 3121 GCTCAGAGCCCTGGAATGAATGTCTCACTGTGCATCATCTTATGTTTACACAGGCC 3180
QY 1061 PheSerGluLeuArgArgAlaGlnMetThrGluGlyProAsnThrAlaProAsnPhe 1080
DB 3181 TTTTCTGAACTTTAGAGTGTCCCAATGACAGAGGCCAAATACAGACCTCCAACTT 3240
QY 1081 SerHisThrGlyProThrPheProValValProPheLeuSerSerIleAlaGlyVal 1100
DB 3241 AGTCATACAGGACCAACATTTCCAGTAGTACTCTTTCTTATAGTACATTTGCTGAGTCT 3300
QY 1101 ProThrThrAlaAlaAlaThrAlaProValProAlaThrSerSerProProAsnAspTlle 1120
DB 3301 CCAACCAACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 3360
QY 1121 SerThrSerValIleGlnSerGluValThrValProThrGlnGluGlyIleAlaGlyVal 1140
DB 3361 TCCATCATCATTAATTCAGTCTGAGTTACAGTCCCTCAAGAGGGAATTCGCGAGTT 3420
QY 1141 AlaThrSerThrGlyValValThrSerGlyGlyLeuProIleProProValSerGluSer 1160
DB 3421 GCCACACGACAGAGTGTGTAACTTCAAGTGTCTCCCAATACCACTGTGTGAAATCA 3480
QY 1161 ProValLeuSerSerValValSerSerTlleThrIleProAlaValAlaSerTlleSerThr 1180
DB 3481 CCAGTACTTTCAGCGTAGTTCAGATACACATACATCTGCAAGTGTCTCAATATCTACT 3540
QY 1181 ThrSerProSerLeuGlnValProThrSerThrSerGlnIleValAlaSerSerThrAla 1200
DB 3541 ACATCCCGTCACTTCAAGTCCCAATCCCATCTCAGATCATGATCTGTGTTCTAGTAGACA 3600
QY 1201 LeuTyrProSerValThrValSerAlaThrSerAlaSerAlaGlySerThrAlaThr 1220
DB 3601 CTGATCTCTCAGTAACAGTTTGCACACTTCACTGAGGGGAGTACGTATACC 3660
QY 1221 ProGlyProLysProProAlaValValSerGlnGlnAlaAlaGlySerThrThrValGly 1240
DB 3661 CCAGTCTCTAAGCTTCCAGCTGTATCTCAGCAGGACAGCAGCAGCAGCAGCAGCAGCAG 3720
QY 1241 AlaThrLeuThrSerValSerThrThrThrSerPheProSerThrAlaSerGlnLeuSer 1260
DB 3721 GCCACATTAACATAGATTTCTACACACATCTTCCAGACACAGCTTCCACACTGTCC 3780
QY 1261 IleGlnLeuSerSerSerThrThrThrProThrLeuAlaGluThrValValSerAla 1280
DB 3781 ATTACAGTTACACAGACTTCTACTCTTACTTACTTACTTACTTACTTACTTACTTACT 3840
QY 1281 HisSerLeuAspLysThrSerHisSerSerThrThrGlyLeuAlaPheSerLeuSerAla 1300
DB 3841 CACTCAGTACAGTACATCTTCAAGCAGTCACTGATGAGTGTGCTTCTCTCTCTCTCTCA 3900
QY 1301 ProSerSerSerSerSerProGlyAlaGlyValSerSerTyrIleSerGlnProGlyIly 1320
DB 3901 CCATCTTCTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3960
QY 1321 LeuHisProLeuValIleProSerValIleAlaSerThrProIleLeuProGlnAlaAla 1340
DB 3961 CTGATCTCTTGTGATTCATTCATCAGTATGATCTTACTCTTACTTCTCTCTCTCTCTCT 4020
QY 1341 GlyProThrSerThrProLeuLeuProGlnValProSerTlleProProLeuValGlnPro 1360

Db 4021 GGACCTACTTCTACACCTTTATTATACCCCAAGTACCTAGTACCACTTGGTGACAGCCT 4080
QY 1361 ValAlaAsnValProAlaValGlnGlnThrLeuIleHisSerGlnProGlnProAlaLeu 1380
Db 4081 GTTGCCAAATGGCTGCTGTAAGACAGACAACTAATTCATATGACCTTCACCACTTGTG 4140
QY 1381 LeuProAsnGlnProHisThrHisCysProGluValAspSerAspThrGlnProHisAla 1400
Db 4141 CTGCCAACAGCCCACTATCTCATTTGTCTCTGAAGTAAATTCCTGATACACACCCAAAGCT 4200
QY 1401 ProGlyIleAspAspIleLysThrLeuGlnGlnLysLeuArgSerLeuPheSerGlnHis 1420
Db 4201 CCTGAATATGATGACATAAAGCTCTAGAGAAAAGCTGCGGTCTCTGTTCAGTGAACAC 4260
QY 1421 SerSerSerGlyAlaGlnHisAlaSerValSerLeuGlnThrSerLeuValIleGlnSer 1440
Db 4261 AGCTATCTGGAGCTCAGCATGCTCTCTGCTCTGACCTGAGAACCTCACTGATACAGAGCC 4320
QY 1441 ThrValThrProGlyIleProThrThrAlaValAlaProSerLysLeuLeuThrSerThr 1460
Db 4321 ACTGTCAACACAGGATCTCCAACTACTGCTGTGACCAAGAAACCTCTGACTTCTAC 4380
QY 1461 ThrSerThrCysLeuProProThrAsnLeuProLeuGlyThrValAlaLeuProValThr 1480
Db 4381 ACAAGTACTTCTTACCAACCAACCAATTTACCACTAGGACAGATTGCTTGCCAGTTACA 4440
QY 1481 ProValValThrProGlyGlnValSerThrProValSerThrThrThrSerGlyValLys 1500
Db 4441 CCAGTGTGACACCTGGGCAAGTTCTTACCCAGTCAACCTACTACATCAAGAGGTGAAA 4500
QY 1501 ProGlyThrAlaProSerLysProProLeuThrLysAlaProValLeuProValGlyThr 1520
Db 4501 CCTGAATCTGCTCCCTCCAGCCACCTCTACTAAGGCTCGGAGTCCGACCTGGTACT 4560
QY 1521 GlnLeuProAlaGlyThrLeuProSerGlnGlnLeuProProPheProGlyProSerLeu 1540
Db 4561 GAACCTTCAGAGGATCTTACCCAGGACGAGCTGCGCACCTTTTCCAGGACCTTCTCTA 4620
QY 1541 ThrGlnSerGlnGlnProLeuGlnAspLeuAspAlaGlnLeuArgArgThrLeuSerPro 1560
Db 4621 ACCGAGTCCCAAGCACTCTTGAAGATCTTATGCTCAATTGAGAAAGAACCTTAGTCA 4680
QY 1561 Glu**1LeThrValThrSerAlaValGlyProValSerMetAlaAlaProThrAlaIle 1580
Db 4681 GAGATKATCAAGAGACTTCTGCGGTGCTGCTGTGTCATGCGGCTCCAAAGCAATC 4740
QY 1581 ThrGlnAlaGlyThrGlnProGlnLysGlyValSerGlnValLysGlnGlyProValLeu 1600
Db 4741 ACAGAAACAGAAACAGCCCTCAGAAAGGTTCTTCAAGTCAAAAGAGGCCCTGCTCTA 4800
QY 1601 AlaThrSerSerGlyAlaGlyValPheLysMetGlyArgPheGlnValSerValAlaAla 1620
Db 4801 GCACATGATTCAGAGAGCTGGGTGTTTAAATGGAAGCATTTTCAAGTTTCTGTTGACGA 4860
QY 1621 AspGlyValaGlnLysGlnGlyLysAsnLysSerGlnAspAlaLysSerValHisPheGln 1640
Db 4861 GACGGTCCCAAGAAAGGCTTAAATAATAGTCAGAAAGATGCAAGTCTGTTTGA 4920
QY 1641 SerSerThrSerGlnSerSerValLeuSerSerSerProGlnSerThrLeuValLys 1660
Db 4921 TCCAGGACCTCAGAGTCTCAGTGTCTATCAAGTAAAGTCCAGAGATACCTTGTTGTA 4980
QY 1661 ProGlnProAsnGlyIleThrIleProGlyIleSerSerAspValProGlnSerAlaHis 1680
Db 4981 CCAGAGCCGAATGATCAATCAATCCCTGGATCTCTTCAATGGCCAGAGAGGCCAC 5040
QY 1681 LysThrThrAlaSerGlnAlaLysSerAspThrGlyGlnProThrLysValGlyArgPhe 1700
Db 5041 AAAAATACTAGCTCTCAGAGGCAAAAGTCAACCTGGGAGCTTACCAAGTGTGACGTTT 5100
QY 1701 GlnValThrThrThrAlaAsnLysValGlyArgPheSerValSerLysThrGlnAspLys 1720
Db 5101 CAGGTGACAACTACACCAAAACAAAGTGGGTCTTCTGTATCAAAAACCTGAGGACAG 5160
QY 1721 IleThrAspThrLysLysGlnGlyProValAlaSerProProPheMetAspLeuGln 1740
Db 5161 ATCACTGACACAAAGAAAGAAAGACCAAGTGCATCTCTCTTATGATTTGACACA 5220
QY 1741 AlaValLeuProAlaValIleProLysLysGlnLysProGlnLeuSerGlnProSerHis 1760
Db 5221 GCTGTTCTTCTGCTGATATACAAAGAAAGAAAGAGCTGTAAGTGTACAGAGCTTCAAT 5280
QY 1761 LeuAsnGlyProSerSerAspProGlnAlaAlaPheLeuSerArgAspValAspAspGly 1780
Db 5281 CTAATAGGCGCGTCTTTCGACCCGAGGCCCTTTTAAATGAGGATGTGATATGCT 5340
QY 1781 SerGlySerProHisSerProHisGlnLeuSerSerLysSerLeuProSerGlnAsnLeu 1800
Db 5341 TCCGATATCCACACTGCCCCATCAGCTGAGTCAAAAGACCTTCTTACGACAGATCTA 5400
QY 1801 SerGlnSerLeuSerAsnSerPheAsnSerSerTyrtMetSerSerAspAsnGlnSerAsp 1820
Db 5401 AGTCAAAAGCTTAGTAAATTCATTTAACTCTTTCATGATGATGAGCAATGATGACAGAT 5460
QY 1821 IleGlnAspGlnAspLeuLysLysGlnLeuArgArgLeuArgAspLysHisLeuLysGln 1840
Db 5461 ATCGAAGATGAAGACTTAAAGTTAGACTGCGAGACTACAGATTAACATCTCAAAAG 5520
QY 1841 IleGlnAspLeuGlnSerArgGlnLysHisGlnIleGlnSerLeuTyrtThrLysLeuGly 1860
Db 5521 ATTGAGACCTGCAAGATGCGCAGAACATGAAATTAATCTTTGTATACCAAACTGGGC 5580
QY 1861 LysValProProAlaValIleIleProProAlaAlaProLeuSerGlyArgArgArgArg 1880
Db 5581 AAGTGTCCCTGCTGTATTTATTTCCCAAGCTGCTCCCTTCAAGGAGAAAGACGA 5640
QY 1881 ProThrLysSerLysGlySerLysSerSerArgSerSerSerLeuGlyAsnLysSerPro 1900
Db 5641 CCCACTAAAGCAAAAGCAGAGAAATCTAGTGAAGAGAGTCTCTGGGAAATMAAGCCCC 5700
QY 1901 GlnLeuSerGlyAsnLeuSerGlyGlnSerAlaAlaSerValLeuHisSProGlnThr 1920
Db 5701 CAGCTTCAGATTAACCTGTCTGTCAGAGTCAAGCTTCAAGTCTTGCACCCCCAGCAAGCC 5760
QY 1921 LeuHisProProGlyAsnIleProGlnSerGlyGlnAsnGlnLeuLeuGlnProLeuLys 1940
Db 5761 CTCCACCTCTCTGGCAACATCCCAAGTCCGGGCAAGATAGCTGTATACGCCCTTAAG 5820
QY 1941 ProSerProSerSerAspAsnLeuTyrtSerAlaPheThrSerAspGlyAlaIleSerVal 1960
Db 5821 CCATCTCCCTCCAGTGAACAACCTCTATTCAGCTTCAACAGTATGCTGCATTTCAAGTA 5880
QY 1961 ProSerLeuSerAlaProGlyGlnGlyThrSerSerThrAsnThrValGlyAlaThrVal 1980
Db 5881 CCAAGCTTCTCTCTCAGGTCAAGGAAACAGAGCAACAAACCTGTTGGGGCAACAGT 5940
QY 1981 AsnSerGlnAlaAlaGlnAlaGlnProProAlaMetThrSerSerArgLysGlyThrPhe 2000
Db 5941 AACAGCCAGCCGCCCAAGTCAAGCTCTCTGCAATGATGCTCCAGCAAGAGAGGCAATTC 6000
QY 2001 ThrAspAspLeuHisLysLeuValAspAsnTrpAlaArgAspAlaMetAsnLeuSerGly 2020
Db 6001 ACAGATGACTTGCACAAAGTTGTAGCAATTTGGGCCGAGAGTCCATGATCTCTAGGC 6060
QY 2021 ArgArgGlySerLysGlyHisMetAsnTyrtGlnGlyProGlyMetAlaArgLysPheSer 2040
Db 6061 AGGAGAGGAAGCAAAAGGACATGAATTAAGAGGCCCTGGAATGCAAGAGATTCCT 6120
QY 2041 AlaProGlyGlnLeuCysIleSerMetThrSerAsnLeuGlyGlySerAlaProIleSer 2060
Db 6121 GCACCTGGGCACTGTGCAATTCATGACCTGCAACCTGGGTGCTCTGCCCCATCTCT 6180
QY 2061 AlaAlaSerAlaThrSerLeuGlyHisPheThrLysSerMetCysProProGlnGlnTyrt 2080
Db 6181 GCAGCATCAGCTACCTCTTAGGTCACTTCAACAAATCTATGAGCCCCCAACAGCAATAT 6240

QY 2081 GlyPheProAlaThrProPheGlyAlaGlnTrpSerGlyThrGlyGlyProAlaProGln 2100
 DB 6241 GGCTTTCAGCTACCCCATTTGGGCTCAATGAGTGAGGAGGAGTGGCCAGCAGCAG 6300
 QY 2101 ProLeuGlyGlnPheGlnProValGlyThrAlaSerLeuGlnAenPheAenLleSerAen 2120
 DB 6301 CCACCTTGCCCACTTCACCACTGCGGAACCTGCTCCCTTCGACGAATTCACATCAGCAAT 6360
 QY 2121 LeuGlnLysSerLleSerAenProGlySerAenLysAenArgThrThr 2136
 DB 6361 TTGCAAAATTCATCAGCAACCCCGAGCTCCACCTCGAGCACT 6408
 RESULT 2
 AAD34306
 ID AAD34306 standard; cDNA; 7280 BP.
 XX
 XX AAD34306;
 XX
 XX 16-JUL-2002 (first entry)
 XX
 XX Human PKIN-9 cDNA.
 DE
 XX Human; kinase; enzyme; PKIN-9 protein; immune system disorder; anaemia;
 KM acquired immune deficiency syndrome; thymic hypoplasia; Crohn's disease;
 KM aschma; neurological disorder; epilepsy; Charcot-Marie-Tooth disease;
 KM AIDS; seizures; cell proliferative disorder; cancer; adenocarcinoma;
 KM leukemia; lymphoma; melanoma; myeloma; sarcoma; developmental disorder;
 KM Down's syndrome; gene therapy; protein therapy; cytostatic; gene; ss.
 OS
 XX Homo sapiens.
 XX
 XX Key Location/Qualifiers
 FH 419..6826
 FT /**tag= a
 FT /product= "Human PKIN-9 protein"
 FT
 XX WO200218557-A2.
 XX
 XX 07-MAR-2002.
 XX
 XX 31-AUG-2001; 2001WO-US027219.
 PF
 XX 31-AUG-2000; 2000US-0229873P.
 PR 08-SEP-2000; 2000US-0231357P.
 PR 14-SEP-2000; 2000US-0232654P.
 PR 22-SEP-2000; 2000US-0234902P.
 PR 29-SEP-2000; 2000US-0236492P.
 PR 06-OCT-2000; 2000US-0238389P.
 PR 13-OCT-2000; 2000US-0240542P.
 XX
 XX (INCY-) INCYTE GENOMICS INC.
 PA
 XX Bandman O, Nguyen DB, Walla NK, Hafalia AJA, Yao MG, Gandhi AR,
 PI Gururajan R, Ding L, Patterson C, Yue H, Baughn WK, Tribouley CM,
 PI Thornton M, Elliott JD, Ison CH, Au-Yang J, Tang YJ,
 PI Azimzai Y, Burdill VS, Marcus GA, Zingler KA, Lu DM, Lal PG,
 PI Rankumar J, Warren BA, Kearney L, Policky UL, Thangavelu K,
 PI Burford N;
 XX
 XX WPI: 2002-329769/36.
 DR P-PSDB; AAE21714.
 DR
 XX New human kinases, useful for diagnosing, treating or preventing immune
 PT system disorders (e.g. Crohn's disease) neurological disorders (e.g.
 PT epilepsy), or cell proliferative disorders (e.g. cancers such as leukemia
 PT or lymphoma).
 XX
 XX Claim 88; Page 203-205; 218pp; English.
 XX
 CC The present invention relates to human kinases (PKIN) and polynucleotides
 CC encoding such proteins. PKIN sequences of the invention are useful for
 CC diagnosing, treating or preventing disorders associated with aberrant
 CC expression of PKIN, particularly immune system disorders (e.g. acquired

CC immune deficiency syndrome (AIDS), thymic hypoplasia, Crohn's disease,
 CC anaemia, aschma), neurological disorders (e.g. epilepsy, Charcot-Marie-
 CC Tooth disease or seizures), cell proliferative disorders (e.g. cancers
 CC such as adenocarcinoma, leukemia, lymphoma, melanoma, myeloma, sarcoma),
 CC and developmental disorders (e.g. Down's syndrome). They are also used in
 CC gene therapy and protein therapy. The present sequence is a cDNA encoding
 CC human PKIN-9 protein
 XX
 SQ Sequence 7280 BP; 1940 A; 1999 C; 1720 G; 1621 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 0 Length: 7280
 Score: 10794.50 Matches: 2134
 Percent Similarity: 99.91% Conservative: 0
 Best Local Similarity: 99.91% Mismatches: 1
 Query Match: 99.84% Indels: 1
 DB: 6 Gaps: 1
 US-10-010-720-14 (1-2136) x AAD34306 (1-7280)
 QY 1 MetSerGlyGlyAlaAlaGluLysGlnSerSerThrProGlySerLeuPheLeuSerPro 20
 DB 419 ATGCTTGGCGCGCCCGCAGAGAGACAGACACTCCGGTCTCTCTCTCGCGG 478
 QY 21 ProAlaProAlaProLysAenGlySerSerSerSerSerValGlyGlyLeuGly 40
 DB 479 CCGGCTCTCGCCCAAGATGCTCCAGCTCCGATTCCTCGGGGAGAAACTGGGA 538
 QY 41 AAlaAlaAlaAlaAlaAlaValAlaThrGlyArgThrGlnGluLysArgArgHisThr 60
 DB 539 GCCCGCGCGCCCGCAGCTGTGACCGGACGACGAGAGATACAGCGCGCGCCACACT 598
 QY 61 MetAspLysAspSerSerArgGlyAlaAlaAlaAlaThrThrThrThrThrGluHisArgPhePhe 80
 DB 599 ATGACAAAGACAGACCGCTGGGGCGCGCCGACCACTACACACTGAGCACTGCTTTC 658
 QY 81 ArgArgSerValLleCysAspSerSerAlaAlaThrAlaLeuGlnLeuProGlyLeuProLeu 100
 DB 659 CGCGGAGCGCTCATCTCGACTCCCATGCACTGAGACTGAGCTTCCGCGCTTCTCTT 718
 QY 101 SerLeuProGlnProSerLleProAlaAlaValAlaProGlnSerAlaProProGluProHis 120
 DB 719 TCCCTGCCCCAGCCACACTCCCGCGCTGTCCGAGAGTGTCTCCACGAGCCCTCAC 778
 QY 121 ArgGluGluThrValAlaThrAlaAlaThrSerGlnValAlaGlnGlnProProAlaAla 140
 DB 779 CGGAAAGACCGTGAACCGCCACCTCCAGGTAGCCGACGACTCCAGCCCT 838
 QY 141 AlaAlaProGlyGlnGlnAlaValAlaGlyProAlaProSerThrValProSerSerThr 160
 DB 839 GCGGCCCTGGGGAACGCGCTCGCGGCGCTCCCTCGACTGTCTCCAGAGTACC 898
 QY 161 SerLysAspArgProValSerGlnProSerLeuValGlySerLysGlnGluProProPro 180
 DB 899 AGCAAAACCGCCAGTGTCCAGCTTGTGGGAGCAAAAGAGAGAGAGAGAGAGAGAG 958
 QY 181 AlaArgSerGlySerGlyGlySerAlaLysGlnProGlnGlnGlnGlnGlnGlnGln 200
 DB 959 GCGAGAGTGGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1018
 QY 201 GlnAspAspLleGluGluLeuGluThrValAlaValGlyMetSerAenAspGlyArgPhe 220
 DB 1019 CAGATGATATCGAAGAGCTGAGAACAGAGCGCTGGGAATGTCTAACGATGCGCTTT 1078
 QY 221 LeuLysPheAspLleGluLleGlyArgGlySerPheLysThrValLysLysGlyLeuAsp 240
 DB 1079 CTCAAGTTGACATCGAATCGAGAGAGCTCTTTAAGACGCTTACAAAGGTGTGAC 1138
 QY 241 ThrGluThrThrValGluValAlaAlaATPcysGluLeuGlnAenArgLysLeuThrLysSer 260
 DB 1139 ACTGAACCAACCGTGAAGAGTGGCTGTGTGATACGAGATTCGAATTAACAAAGCT 1198
 QY 261 GluArgGlnArgPheLysGluGluAlaGluMetLeuLysGlyLeuGlnHisProAsnIle 280

```
Db 1199 GAGAGCAGAGATTAAAGAGAACTGAAATGTTAAAAAGCTTCCAGCATCCAAATATT 1258
Qy 281 ValArgPheTyArgSerSerTyrGluSerThrValLysGlyLysLysCysIleValLeuVal 300
Db 1259 GTTAATATTTTATGATTCCTGGGAAATCCACAGTAAAGAGAAAGAAATGCAATGTTTGGTGG 1318
Qy 301 ThrGluLeuMetThrSerGlyThrLeuLysThrTyrLeuLysArgPheLysValMetLys 320
Db 1319 ACTGAACCTTATGACCTCTGGAAACCTTAAACGTTATCTGAAAAGGTTTAAAGTATGAAAG 1378
Qy 321 IleLysValLeuArgSerSerTyrCysArgGlnIleLeuLysGlyLeuGlnPheLeuHisThr 340
Db 1379 ATCAAGATTCTAAGAGCTGGGCGCTGAGATCTTAAAGGCTTCAGATTCTTCATACT 1438
Qy 341 ArgThrProProlIleLeHisArgAspLeuLysCysAspAsnIlePheIleThrGlyPro 360
Db 1439 CGAACTCCAACTATCATTCACCGCGATCTTAAATGTGACAACTGTTATACCGGCGCTT 1498
Qy 361 ThrGlySerValLysIleGlyAspLeuGlyLeuAlaThrLeuLysArgAlaSerPheAla 380
Db 1499 ACTGGCTCAGTCACAGATTGGAGACCTCGGTCTGGGCAACCTGAAAGCGGCTTCTTTGCC 1558
Qy 381 LysSerValIleGlyThrProGluPheMetAlaProGluMetTyrGluGluLysTyrAsp 400
Db 1559 AAGAGTGTAGTAGTACCCAGAGATTCAAGCCCTCGAGATGTATGAGAGAAATATGAT 1618
Qy 401 GluSerValAspValTyrAlaPheGlyMetCysMetLeuGluMetAlaThrSerGlyLys 420
Db 1619 GAATCCGTTGACGTTATGCTTTGGAGTGTGACATGCTTGAGATGGCTACATCTGAAATAT 1678
Qy 421 ProTyrSerGluCysGlnAsnAlaAlaGlnIleTyrArgArgValThrSerGlyValLys 440
Db 1679 CCTTACTCGAGGTGCAAAAGTCTGGCGCATCTACCTCCCGTGAACCAAGCGGCTGAG 1738
Qy 441 ProAlaSerPheAspLysValAlaIleProGluValLysGlnIleIleGluGlyCysIle 460
Db 1739 CCAGCCAGTTTTCACAAAGTACAAATCTCGAAGAGAAATATATGAGAGATGCTATA 1798
Qy 461 ArgGlnAsnLysAspGluArgTyrSerIleLysAspLeuAsnHisAlaPhePheGln 480
Db 1799 CGAATAACAAAGATGAAGATATCTCCATCAAAAGCCTTGTGAACCAATGCTTCTTCAA 1858
Qy 481 GlnGluThrArgValArgValGluLeuAlaGluLysAspGlyLysIleAlaIle 500
Db 1859 GAGAAAACAGAGATACGGGTAGATTTAGCAGAAAGATGATGAGAAAAATAGCCATA 1918
Qy 501 LysLeuTyrPheLeuArgIleGluAspIleLysLysLeuLysGlyLysTyrLysAspAsnGlu 520
Db 1919 AAATTTATGCGTACGATTTGAAGATTTAAGAAATTTAAAGGAAATATCAAAAGATATGAA 1978
Qy 521 AlaIleGluPheSerPheAspLeuGluArgAspValProGluAspValAlaGlnGluMet 540
Db 1979 GCTATTGAGTTTCTTTGATTATTAAGAGATGTCCTCCAGAAAGATGTTGCCAAAGAAATG 2038
Qy 541 ValGluSerGlyTyrValCysGluGlyAspHisLysThrMetAlaLysAlaIleLysAsp 560
Db 2039 GTAGAGCTCGGGTATGCTGTGAAGGTGATCACAAGACCATGGCTTAAAGCATATCAAGAC 2098
Qy 561 ArgAlaSerLeuLysLysArgLysArgGluLysArgGlnLeuValArgGluGlnGlu 580
Db 2099 AGAATATCATTAATTAAGAGAAACGAGACGCGGAGTTGTATCGGAGAGAACAAAGAA 2158
Qy 581 LysLysLysGlnGluGluSerSerLeuLysGlnIleValGluGlnSerSerAlaSerGln 600
Db 2159 AAAAAAAGCAGAGAGAGACAGTCTCAACACAGAGGTAGAGAACATCCAGTCTTCCAG 2218
Qy 601 ThrGlyLysGlnLeuProSerAlaSerThrGlyIleProThrAlaSerThrThrSer 620
Db 2219 ACAGAGATCAAGACGCTCCCTTCTGCTAGCACCGGCACTACTACTTCTACCACTTCA 2278
Qy 621 AlaSerValSerThrGlnValGluProGluGluProGluAlaAspGlnHisGlnGlnLeu 640
Db 2279 GCTTCAGTTTCTACACAACTGAACTGAAACACTTGAGGACATCAATCAACACTA 2338
Qy 641 GlnTyrGlnGlnProSerIleSerValLeuSerAspGlyThrValAspSerGlyGlnGly 660
Db 2339 CAGTACCAAGCAACCCAGATATATCTGTATTATCTGATGAGCGGTTGACAGTGTACAGGA 2398
Qy 661 SerSerValPheThrGluSerArgValSerSerGlnIleThrValSerTyrGlySerGln 680
Db 2399 TCTCTGCTTCTCAACAATCTCGAGTGTGAGCCACCAAGACAGATTCATATAGTTCCTCAA 2458
Qy 681 HisGlnGlnAlaHisSerThrGlyThrValProGlyHisIleProSerThrValGlnAla 700
Db 2459 CATGAACAGGACATTTACAGGACAGATCCCAAGGCAATATACCTTACTACTCCAAACA 2518
Qy 701 GlnSerGlnProHisGlyValTyrProProSerSerValAlaGlnGlyGlnSerGlnGly 720
Db 2519 CAGTCTCAGCCCATGGGGTATATCCACCCCAAGTGTGGACAGGGGCGACAGCGAGGT 2578
Qy 721 GlnProSerSerSerSerLeuThrGlyValSerSerSerGlnProIleGlnHisProGln 740
Db 2579 CAGCCATCTCAAGTACGTTTACAGGGGTTTCAATCTTCCAAACCATACATCAATCT--- 2635
Qy 741 GlnGlnGlnGlyIleGlnIleThrAlaProProGlnGlnThrValGlnTyrSerLeuSer 760
Db 2636 CAGCAGCAGGAAATACAGACAGACAGCCCTCCTCAACAGACAGTGCATTTCACTTTCA 2695
Qy 761 GlnThrSerThrSerSerGluAlaThrThrAlaGlnProValSerGlnProGlnAlaPro 780
Db 2696 CAGACATCAACCTCAGTGAAGCCATCTGACACAGCCAGTGAATCAACCTCAAGCTCCA 2755
Qy 781 GlnValLeuProGlnValSerAlaGlyLysGlnSerThrGlnGlyValSerGlyValAla 800
Db 2756 CAAGTCTTGCTTCAAGTATCACTGAGAAACAGAGTACTCAGGAGTCTTCAAGGTGCT 2815
Qy 801 ProAlaGluProValAlaValAlaGlnProGlnAlaThrGlnProThrThrLeuAlaSer 820
Db 2816 CCTCAGACAGCCAGTGCAGTGCAGACAGCCCAAGCTCCAGCCGACCACTTGGCTTCC 2875
Qy 821 SerValAspSerAlaHisSerAspValAlaSerGlyMetSerAspGlyAsnGluAsnVal 840
Db 2876 TCTGTAGACAGTGCACATTCAGATGTTCTTCAAGTATGATGTATGCAATGACAAAGTCT 2935
Qy 841 ProSerSerSerGlyArgHisGluGlyArgThrThrLysArgHisTyrArgLysSerVal 860
Db 2936 CCATCTTCCATGGAAGGACATGAAGAGAAAGTACAAACGCGCTTACCGAAATCTGTA 2995
Qy 861 ArgSerArgSerArgHisGluLysThrSerArgProLysLeuArgIleLeuAsnValSer 880
Db 2996 AGAGTGTGCTCTGCATGAATAAACTTCAAGCCCAAAATTTAAGATTTTGAATGTTTCA 3055
Qy 881 AsnLysGlyAspArgValAlaGluCysGlnLeuGlnIleThrHisAsnArgLysMetValThr 900
Db 3056 AATTAAGAGAGCCGAGTAGATGAATGTCAATTTAGAGACTCAATATATGAAATATGTTACA 3115
Qy 901 PheLysPheAspLeuAspGlyAspAspProGlnIleAlaThrIleMetValAsnAsn 920
Db 3116 TTCAAAATTTGACCTAGATGATGACAAACCCGAGAGATGACCAACAATTTATGTTAAACAT 3175
Qy 921 AspPheIleLeuAlaIleGluArgGluSerPheValAspGlnValArgGluIleIleGlu 940
Db 3176 GACTTTATTTCTAGCAATAGAGAGAGAGTCTGTTGTGATCAAGTGCAGAAATATATGAA 3235
Qy 941 LysAlaAspGlnMetLeuSerGluAspValSerValGlnProGluGluLysAspGlnGlyLeu 960
Db 3236 AAAGCTGATGAATATGCTCACTGAGAGATGTCAGTGTGAACCAAGAGGTGATACAGGAGTTG 3295
Qy 961 GluSerLeuGlnGlnLysAspAspTyrGlyPheSerGlySerGlnLysLeuGluGlyGlu 980
Db 3296 GAGAGTCTACAGAAAGATGACTATGCTTTCAAGTCTCAGAAATATGAAAGAGAG 3355
Qy 981 PheLysGlnProIleProAlaSerSerMetProGlnGlnIleGlyIleProThrSerSer 1000
Db 3356 TTCAAACAAACCAATTCCTGCGTCTTCCATGCAACAGCAAAATAGCATTCCTACAGTTCT 3415
```

QY 1001 LeuThrGlnValAlaHisSerAlaGlyArgPheIleValSerProValProGlnSer 1020
Db 3416 TTAACCTAGGTGTTCTTCGCGGGAAGGGTTTATGAGTCTCTGCGCAAGAAC 3475
QY 1021 ArgLeuArgGlnSerLysValPheProSerGlnIleThrAspThrValAlaSerThr 1040
Db 3476 CGATTACAGAACATCAAAAGTTTCCCGAGTAATAACAGATACAGTTGGCTGCTTACA 3535
QY 1041 AlaGlnSerProGlyMetLeuSerHisSerAlaSerSerLeuSerGlnGlnAla 1060
Db 3536 GCTCAGAGCCCTGGAAAGAACTTCTCAGCTGCAATCCTTAGCTTACCAACAGACC 3595
QY 1061 PheSerGlnLeuArgAlaGlnMetThrGlnGlyProAsnThrAlaProProAsnPhe 1080
Db 3596 TTTTCTGACTTACACGTGGCCAAATGACAGAAAGCAATACAGACCTCCAAACTTT 3655
QY 1081 SerHisThrGlyProThrPheProValValProProPheLeuSerIleAlaGlyVal 1100
Db 3656 AGTCATACAGACCAACATTTCCAGTAGTACCTCTTCTTAAGTACATTTGCTGGAGTC 3715
QY 1101 ProThrThrAlaAlaIleThrAlaProValProAlaThrSerSerProProAsnIle 1120
Db 3716 CCAACACAGCAGCAGCAGCAGCAGCTCTGCAACAGACCCCTCTTAATGACATT 3775
QY 1121 SerThrSerValIleGlnSerGlnValThrValProThrGlnGlnIleAlaGlyVal 1140
Db 3776 TCCACATCAGTAATTCAGTCTGAGTACAGTCCCACTGAAAGAGGAGTTGCTGAGTT 3835
QY 1141 AlaThrSerThrGlyValValThrSerGlyGlyLeuProIleProValSerGlnSer 1160
Db 3836 GCCACCCAGCAGAGTGTGTAACTTCAGTGTCTCCCAACCACTGTGTCTGAATCA 3895
QY 1161 ProValLeuSerSerValIleSerSerIleThrIleProAlaValIleSerIleSerThr 1180
Db 3896 CCACTACTTTCAGCGAGTTCAGATACCAATACCTGAGTGTCTCAATATCACT 3955
QY 1181 ThrSerProSerLeuGlnValProThrSerThrSerGlnIleValIleSerSerThrAla 1200
Db 3956 ACATCCCGCTCACTTCAAGTCCCAACATCCACATCGATCGTTGTTCTAGTACAGCA 4015
QY 1201 LeuThrProSerValIleThrValSerAlaThrSerAlaSerAlaGlySerThrAlaThr 1220
Db 4016 CTGTATCTTTCAGTACAGTTTCAGCAACTTCAGCTCTGAGGGGAGACTGCTTACC 4075
QY 1221 ProGlyProLysProProAlaValIleSerGlnIleAlaGlySerThrThrValGly 1240
Db 4076 CCAAGTCTTAAGCTTCAAGCTGATATCTCAGACGACGACGACACTACTGTGAG 4135
QY 1241 AlaThrLeuThrSerValSerThrThrThrSerPheProSerThrAlaSerGlnLeuSer 1260
Db 4136 GCCACATTAACATCAGTTTCTACACACACTTCATTCCAAGCAGACAGCTCACTGTCC 4195
QY 1261 IleGlnLeuSerSerSerThrThrProThrLeuAlaGlnThrValValIleSerAla 1280
Db 4196 ATTACGCTTACACAGACTTCTACTCTTCTTAGCTGAACCGTGTGTAGTACGCCA 4255
QY 1281 HisSerLeuAspLysThrSerHisSerSerThrThrGlyLeuAlaPheSerLeuSerAla 1300
Db 4256 CACTCAGTACATGACATCTCATACGACGATCAACTGATGGCTTCTCTCTCTGCA 4315
QY 1301 ProSerSerSerSerSerProGlyAlaGlyValSerSerThrIleSerGlnProGlyGly 1320
Db 4316 CCACTTCTCTCTCTCTCTCTGAGGAGAGAGTGTCTAGTTATATTCTCAGCCCTGTGG 4375
QY 1321 LeuHisProLeuValIleProSerValIleAlaSerThrProIleLeuProGlnAlaAla 1340
Db 4376 CTGATCTCTTGTGCTATTCATGATGATGATGATGATGATGATGATGATGATGATGAT 4435
QY 1341 GlyProThrSerThrProLeuLeuProGlnValProSerIleProProLeuValGlyPro 1360
Db 4436 GGACCTACTTCTTACACCTTATATACCCCAAGTACCTAGTATACCACTTGTGTACAGCT 4495
QY 1361 ValAlaAsnValProAlaValGlnGlnThrLeuIleHisSerGlnProGlnProAlaLeu 1380
Db 4496 GTTGCCAAATGTGCTGCTGTACAGACAGACTAATCATATGATGAGCTCAACAGCTTTG 4555
QY 1381 LeuProAsnGlnProHisThrHisGlySerProGlnValAspSerAspThrGlnProValAla 1400
Db 4556 CTTCACCAACAGCCCACTACTCTATGCTCCGAAAGTATGATCTATATACCAACCAAGCT 4615
QY 1401 ProGlyIleAspAspIleLeuThrLeuGlnGlyLeuSerLeuPheSerGlnHis 1420
Db 4616 CTTGGAATTTATACATTAAGACTTGAAGAAAGAGTGGGTCTCTGTTCAAGGAACAC 4675
QY 1421 SerSerSerGlyAlaGlnHisAlaSerValSerLeuGlnThrSerLeuValIleGlnSer 1440
Db 4676 AGCTACTGTGAGTGCATGCTCTGTCTCACTGAGACCTCACTAGTATGAGAGC 4735
QY 1441 ThrValThrProGlyIleProThrThrAlaValAlaProSerLysLeuLeuThrSerThr 1460
Db 4736 ACTGTACACAGAGCACTCCACTACTGTGTGACCAAGCAAACTCTGACTTCTTACC 4795
QY 1461 ThrSerThrCysLeuProProThrAsnLeuProLeuGlyThrValAlaLeuProValThr 1480
Db 4796 ACNAGTACTGCTTACCAACCAACATTTACCTAGAAACAGTGTCTTGCCAGTTACA 4855
QY 1481 ProValValThrProGlyGlnValSerThrProValSerThrThrThrSerGlyValLys 1500
Db 4856 CCACTGTGTACACCTGGGCAAGTTTCTACCCAGTCAAGCTACTACATCAGAGAGTAAA 4915
QY 1501 ProGlyThrAlaProSerLysProProLeuThrLysAlaProValLeuProValGlyThr 1520
Db 4916 CTTGGAATCTGCTCTTCAAGCACTCTAAGCTCTGAGTCTGAGTCTGAGTCTGAGTCT 4975
QY 1521 GlnLeuProAlaGlyThrLeuProSerGlnGlnLeuProProPheProGlyProSerLeu 1540
Db 4976 GAATCTTCAAGAGTACTTACCAAGGACAGAGTGCACCTTTTCAAGACCTTCTCTA 5035
QY 1541 ThrGlnSerGlnGlnProLeuGlnAspLeuAspAlaGlnLeuArgThrLeuSerPro 1560
Db 5036 ACCAGTCCAGCAACTCTTACAGATCTTGAAGTCTCAATTGAGAAACACTTATGTCCA 5095
QY 1561 Gln***IleThrValThrSerAlaValGlyProValSerMetAlaAlaProThrAlaIle 1580
Db 5096 GAGATGATCAGAGTCTGCTGCTGTGTCTGTGTCTGTGTCTGTGTCTGTGTCTGTGTCT 5155
QY 1581 ThrGlnAlaGlyThrGlnProGlnLysGlyValSerGlnValLysGlnGlyProValLeu 1600
Db 5156 ACAGAAAGCAAGACAGCTCTCAGAGAGGAGTCTTCAAGTCAAGAAAGAGCCCTGTCTA 5215
QY 1601 AlaThrSerSerGlyAlaGlyValPheLysMetGlyArgPheGlnValSerValAlaAla 1620
Db 5216 GCAACTAGTTCAGGAGCTGTGTTTAAATGGAGATTCAGTTCAGTTCAGTTCAGTTCAG 5275
QY 1621 AspGlyAlaGlnLysGlnGlyLysAsnLysSerGlnAspAlaLysSerValHisPheGln 1640
Db 5276 GACGCTGCCAGAAAGGGGTAAATAATATCATGAAATCAAAAGTCTGTCAATTGAA 5335
QY 1641 SerSerThrSerGlnSerSerValLeuSerSerSerSerProGlnSerThrLeuValLys 1660
Db 5336 TCCAGACCTCAGAGTCTCTAGTCTTCAAGTATGATCCAGAGATCACTTGTGTA 5395
QY 1661 ProGlnProAsnGlyIleThrIleProGlyIleSerSerAspValProGlnSerAlaHis 1680
Db 5396 CCAAGGCCGAATGACATACCATCTCCGATATCTTCAAGATGTGCCAGAGAGTCCAC 5455
QY 1681 LysThrThrAlaSerGlnAlaLysSerAspThrGlyGlnProThrLysValArgPhe 1700
Db 5456 AAAACTACTGCTTACAGGCAAGTCAAGACCTGGAGGCTTACCAAGAGTGTGGCGTTT 5515
QY 1701 GlnValThrThrThrAlaAsnLysValGlyArgPheSerValSerLysThrGlnAspLys 1720
Db 5516 CAGGTGACACTACAGCAAAAGTGTGTCTGTCTGTATCAAAAACTGAGAGACAG 5575
QY 1721 IleThrAspThrLysLysGlnGlyProValAlaSerProProPheMetAspLeuGln 1740


```

Db 5576 ATCACTGACACAAAGAAAGAGACCAAGGACATCTCCCTTTATGGATTGGACA 5635
Qy 1741 AAlaValLeuProAlaValIleProLysylSerGlyLeuSerGluProSerHis 1760
Db 5636 GCTGTTCTCTCCGCTGTAATCAAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 5695
Qy 1761 LeuAsnGlyProSerSerAspProGluAlaAlaPheLeuSerArgAspValAspAspGly 1790
Db 5696 CTAAATGGGCGCTCTCTGACCCCGAGCGCTTTTAAAGTAGGAGATGATGATGAT 5755
Qy 1781 SerGlySerProHisSerProHisGlnLeuSerSerLysSerLeuProSerGlnAsnLeu 1800
Db 5756 TCCGGTAGTCCACACTCCGCCCATCAGCTGAGCTCAAAAGAGCTTCCTAGCAGATCTA 5815
Qy 1801 SerGlnSerLeuSerAsnSerPheAsnSerSerLysSerSerSerSerAspAsnGluSer 1820
Db 5816 AGTCAAAGCTTAGTAATTCATTTAATCTCTTACATGAGTAGGCAATGAGTCAAGAT 5875
Qy 1821 IleGluAspGluAspLeuLysLeuGluLeuArgArgLeuArgAspLysHisLeuLysGlu 1840
Db 5876 ATCGAAGATGAAGACTTAAGTTAGAGCTGCGACGACATCAAGATTAACATCTCAAAG 5935
Qy 1841 IleGluAspLeuGlnSerArgGlnLysHisGlnIleGluSerLeuLysThrLysLeuGly 1860
Db 5936 ATTCAGAGCTGCGAGAGTCCGCGAAGCATGAATGAATCTTGTATACCAAACTGAGC 5995
Qy 1861 LysValProProAlaValIleIleProProAlaAlaProLysSerGlyArgArgArg 1880
Db 5996 AAGGTGCCCCCTGCTGTTATTAATTCCTCCAGCTGCTCCCTTTCAAGGAGAAAGCA 6055
Qy 1881 ProThrLysSerLysGlySerLysSerSerSerSerSerSerSerLysGlyAsnLysSer 1900
Db 6056 CCCACTAAAGCAAGAGAGAGCAAGATCTAGTCAGAGCACTTCTTGGGAAATTAAGCCCC 6115
Qy 1901 GlnLeuSerGlyAsnLeuSerGlyGlnSerAlaAlaSerValLeuHisProGlnGlnThr 1920
Db 6116 CAGCTTCAGGTAACCTGCTGTCAGAGTCAAGCTTCACTTGCACCCCGAGAGACC 6175
Qy 1921 LeuHisProProGlyAsnIleProGlnSerGlyGlnAsnGlnLeuGlnProLeuLys 1940
Db 6176 CTCAACCTCTCGGACATCCAGAGTCCGAGGAGATCGCTGTAACACCCCTTAAAG 6235
Qy 1941 ProSerProSerSerAspAsnLeuLysSerAlaPheThrSerAspGlyAlaIleSerVal 1960
Db 6236 CCATCTCCCTCCAGTGAACACTCTATTCACTTCACTGATGATGGCTTTCAGTA 6295
Qy 1961 ProSerLeuSerAlaProGlyGlnGlyThrSerSerThrAsnThrValGlyAlaThrVal 1980
Db 6296 CCAAGCCTTCTGCTCCAGTCAAGTCAAGAAACAGAGCAACACTGTTGGGGCAACG 6355
Qy 1981 AsnSerGlnAlaAlaGlnAlaGlnProProAlaMetThrSerSerArgLysGlyThrPhe 2000
Db 6356 AACAGCCAGAGCCGCGCAAGCTGAGCTCTCTCCATGAGTCCACAGGAAAGGAGCA 6415
Qy 2001 ThrAspAspLeuHisLysLeuValAspAsnTrpAlaArgAspAlaMetAsnLeuSerGly 2020
Db 6416 ACAATATGACTTGCAAGTGTGTAGACATTTGGGCCGAGATGCCATGATCTCTCAGGC 6475
Qy 2021 ArgArgLysSerLysGlyHisMetAsnTrpGlnGlyProGlyMetAlaArgLysPheSer 2040
Db 6476 AGGAGAGAGAGCAAGAGGACATGATATGAGGCCCTCGAATATGCAAGAGAAATCTCT 6535
Qy 2041 AlaProGlyGlnLeuCysLysSerMetThrSerAsnLeuGlyGlySerAlaProIleSer 2060
Db 6536 GCACCTGGGCACTGTGATCTTCATGACTTCGAAACCTGGGTGGCTCTGCCCATCTCT 6595
Qy 2061 AlaAlaSerAlaThrSerLeuGlyHisPheThrLysSerMetCysProGlnGlnLys 2080
Db 6596 GCAGCATCAGTACTCTCTAGGTCACTTCAACAGTCTATGTGCCCCCAGACAGATAT 6655
Qy 2081 GlyPheProAlaThrProPheGlyAlaGlnTrpSerGlyThrGlyGlyProAlaProGln 2100

```

```

Db 6656 GGCTTCCAGCTACCCCATTTGGGCTCATATGAGTGGGACGGTGGCCAGCACACAG 6715
Qy 2101 ProLeuGlyGlnPheGlnProValGlyThrAlaSerLeuGlnAsnPheAsnLysSerAsn 2120
Db 6716 CCAGTTGGCCAGTTCCAACTGCTGGGAACTGCTCTCTGAGAAATTTCAATCATCAGCAAT 6775
Qy 2121 LeuGlnLysSerLysSerAsnProProGlySerAsnLeuArgThrThr 2136
Db 6776 TTGCAAAATTCATGCAACCCCGAGCTCCAACTGCGGACCACT 6823

RESULT 3
ID ACH03767
XX ACH03767 standard; DNA; 6690 BP.
AC ACH03767;
XX 26-SEP-2003 (first entry)
DT DNA encoding novel human protein kinase #4.
DE Human; kinase; ds; gene; cosmetic application; nutraceutical application.
XX Homo sapiens.
XX US6541252-B1.
XX 01-APR-2003.
XX 14-MAY-2001; 2001US-00854856.
XX 19-MAY-2000; 2000US-0206015P.
XX (LEXI-) LEXICON GENETICS INC.
XX Walke DW, Hilbun E, Donoho G, Turner CA;
XX WPI; 2003-575927/54.
XX P-PSDB; ABO44384.
XX New nucleic acid encoding novel human proteins, useful in cosmetic and
XX nutraceutical applications.
XX Disclosure; Page; 11pp; English.
XX The invention relates to a new isolated nucleic acid encoding a novel
XX human protein kinase. The nucleic acid is useful in cosmetic and
XX nutraceutical applications. The present sequence represents DNA encoding
XX a novel human protein kinase. Note: The sequence data for this patent did
XX not form part of the printed specification but was obtained in electronic
XX format directly from USPO at
XX seqdata.uspto.gov/sequence.html?docid=6541252B1
XX Sequence 6690 BP; 1842 A; 1808 C; 1532 G; 1505 T; 0 U; 3 Other;

Alignment Scores:
Pred. No.: 0 Length: 6690
Score: 10753.50 Matches: 2136
Percent Similarity: 95.83% Conservative: 0
Best Local Similarity: 95.83% Mismatches: 0
Query Match: 99.46% Indels: 93
Db: 8 Gaps: 1

US-10-010-720-14 (1-2136) x ACH03767 (1-6690)
Qy 1 MetSerGlyGlyAlaAlaGlnLysGlnSerSerThrProGlySerLeuPheLeuSerPro 20
Db 1 ATGTCTGCGCGCGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 60
Qy 21 ProAlaProAlaProLysAsnGlySerSerSerSerSerSerValGlyGlyLysLeuGly 40
Db 61 CCGGCTCTCTCCCGCAAGATGGCTCCAGCTCCGATTCCTCCGCGGAGAAACTGGGA 120
Qy 41 AAlaAlaAlaAspAlaValThrGlyArgThrGluGluLysArgArgArgArgHisThr 60

```



```
QY 781 GlnValIleuProGlnValSerAlaGlyLysGln----- 791
Db 2341 CAAGCTTGCCCTCAAGATCAAGCTGAAAACAGGGCTTCCCACTCGACTGCCACCACAG 2400
QY 791 ----- 791
Db 2401 TAACCAAGAGATTCAATATGTCCTCCCTCTTCCAAAGTGGCTTGTGTTCATTCATCT 2460
QY 791 ----- 791
Db 2461 ACAATCTTAAWCCCTCCCTCCATGCCAGAGAAATGAGCTGACCTGGGTAATCTTCCCA 2520
QY 791 ----- 791
Db 2521 GTGGTGCAGCCTTATGTGAATCAATCTTTAGTTCTTATGGGTGGTGTAGAGAGACAG 2580
QY 791 ----- 791
Db 2581 GTTCAAGTGTCCCAAGCAGAGAGATTAGCACAAGCCCACTACATCTCCACAGCA 2640
QY 792 -----SetThrGlnGlyValSerGlnValAlaProAlaGluProValAlaVal 807
Db 2641 GCAGTTTGGAGAGTACTCAGGAGTCTCTCAGGTTGCTCTGCAAGCCAGTTGCAGTA 2700
QY 808 AlaGlnProGlnAlaThrGlnProThrThrLeuAlaSerSerValAspSerAlaHisSer 827
Db 2701 GCACAGCCCCAAGCTACCCAGCCAGCACTTGGCTTCTCTGTAGACAGTGCACATTTCA 2760
QY 828 AspValAlaSerGlyMetSerAspGlyAsnGluAsnValProSerSerSerGlyValHis 847
Db 2761 GATGTGCTTACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2820
QY 848 GluGlyAlaGlyThrThrLysArgHisTyrArgLysSerValArgSerArgHisGlu 867
Db 2821 GAAGGAATACTACAAACCGCATTTACCGAAATCTGTAGAGAGTGGCTTCGACATGAA 2880
QY 868 LysThrSerArgProGlyLysLeuArgLysLeuAsnValSerAsnLysGlyAspArgVal 887
Db 2881 AAACTTCACGCCCAATTAAGAAATTTGAATGTTCAATTAAGAAAGAGACCGAGTAGTA 2940
QY 888 GluGlyGlnLeuGluThrHisAsnArgLysMetValThrPheLysPheAspLysGly 907
Db 2941 GAATGTCAATTAGAGACTCATTAATAGAAATGTGTACATTCATTAATGACCTGATGT 3000
QY 908 AspAsnProGlnGluIleAlaThrIleMetValAsnAsnAspPheIleLeuAlaIleGlu 927
Db 3001 GACAAACCCCGAGAGATGCAACAATTATGTGAAACAATGACTTATTTCTAGCAATAGAG 3060
QY 928 ArgGlnSerPheValAspGlnValArgGluIleIleGlnLysAlaAspGluMetLeuSer 947
Db 3061 AGAGAGTCTGTTGTGATCAAGTGCAGAAATTTATGAAAAGCTGATGAATGCTCAGT 3120
QY 948 GluAspValSerValGluProGlnGlyAspGlnGlyLeuGlnSerLeuGlnGlyLysAsp 967
Db 3121 GAGGATGTCAATGTGGAAACCAAGAGGTGATCAGGATGTGAGAGTCTCAACAAGAAAGAT 3180
QY 968 AspTyrGlyPheSerGlySerGlnLysLeuGlnGlyGluPheLysGlnProIleProAla 987
Db 3181 GACTATGGCTTTTCAAGTCTCAACAATTTGAAAGAGATTCACAACAACCAATTCCTGCG 3240
QY 988 SerSerMetProGlnGlnIleGlyLysProThrSerSerLeuThrGlnValAlaHisSer 1007
Db 3241 TCTTCATGCGACAGCAAAATGAGCATTCCTTCAGCTTTTAACTCAAGTTGTTCAATCT 3300
QY 1008 AlaGlyArgArgPheIleValSerProValProGlnSerArgLeuArgLysSerVal 1027
Db 3301 GCGGGAAGGCGGTTTATAGTAGTCTGTGCGAAGAGCCCATTAACAAGATCAAAAGTT 3360
QY 1028 PheProSerGluIleThrAspThrValAlaAlaSerThrAlaGlnSerProGlyMetAsn 1047
Db 3361 TTCCCGAGTAATTAACAGATACAGTTGCTGCTCTCAAGCTCAGAGCCCTGAGATGAAC 3420

QY 1048 LeuSerHisSerAlaSerSerLeuSerLeuGlnGlnAlaPheSerGluLeuArgArgAla 1067
Db 3421 TTGTCTCACTGCAATCAATCCCTTAGTGTACAAAGGCGCTTTTCTGAACCTTAGAGCTGC 3480
QY 1068 GlnMetThrGlnGlyProAsnThrAlaProProAsnPheSerHisThrGlyProThrPhe 1087
Db 3481 CAATGACAGAAAGAGCCCAAAVACAGCACTTCCAACTTATGATCAACAGACCAATTT 3540
QY 1088 ProValValProProPheLeuSerSerIleAlaGlyValProThrThrAlaAlaIleThr 1107
Db 3541 CCAGTAGTACCTCTCTTCTTAAAGTAGCATGTGAGTCCCAACACAGCAGCAGCACA 3600
QY 1108 AlaProValProAlaThrSerSerProProAsnIleSerThrSerValIleGlnSer 1127
Db 3601 GCACCACTCCCTGCAACAGAGAGCCCTCTTAATGACATTTCCATCACTGATTAATGACT 3660
QY 1128 GluValThrValProThrGlnGlnGlyIleAlaGlyValAlaThrSerThrGlyValVal 1147
Db 3661 GAGGTTCAGTGCACCACTGAAAGAGGGATGCTGGAGTTGCCACAGCAGAGTGTGTA 3720
QY 1148 ThrSerGlyGlyLeuProIleProProValSerGluSerProValLeuSerSerValVal 1167
Db 3721 ACTTCAGGTGATCTCCCATACACCTGTGTGATCAATCACAGTACTTCCAGCGTAGTT 3780
QY 1168 SerSerIleThrIleProAlaValValSerIleSerThrThrSerProSerLeuGlnVal 1187
Db 3781 TCAGTATCAACAAATCTGCGAGTGTGTCTCAATATCTACATCCCGCTCACTTAAGTC 3840
QY 1188 ProThrSerThrSerGlnIleValValSerSerThrAlaLeuTyrProSerValThrVal 1207
Db 3841 CCCACATCAATCATCTGAGATGTTGTTCTTAGTACAGCAGTATCTTCACTGTAACATCT 3900
QY 1208 SerAlaThrSerAlaSerAlaGlyGlySerThrAlaThrProGlyProLysProProAla 1227
Db 3901 TCACCAACTTACAGCTCTGCGAGGAGGAGTACTGTCTACCCAGGATCTTCAAGCTCCAGCT 3960
QY 1228 ValValSerGlnGlnAlaAlaGlySerThrThrValGlyAlaThrLeuThrSerValSer 1247
Db 3961 GTAGTATCTCGACAGCAGCAGCAGCAGCATCTGTGGAGGCCATTAATCACTCAGTTCT 4020
QY 1248 ThrThrThrSerPheProSerThrAlaSerGlnLeuSerIleGlnLeuSerSerSerThr 1267
Db 4021 ACCACCACTTCATTTCCCAAGCAGCTTCCACAGCTGTCATTCAGCTTAGCAGCAGTACT 4080
QY 1268 SerThrProThrIleAlaGluThrValValValSerAlaHisSerLeuAspLysThrSer 1287
Db 4081 TCTACTCTTACTTAGCTGAAACCGTGTGATGAGCCACACTCAAGATTAAGACATCT 4140
QY 1288 HisSerSerThrThrGlyLeuAlaPheSerLeuSerAlaProSerSerSerSerPro 1307
Db 4141 CATAGCAGTACAACTGATTTGGCTTTCTCTCTCTGACACCATTTCTTCTCTCTCTCT 4200
QY 1308 GlyAlaGlyValSerSerTyrIleSerGlnProGlyGlyLeuHisProLeuValIlePro 1327
Db 4201 GAGAGCAGAGGTGTATGTTATTTCTCAGCTGTGGGTGATCTCTTGTGCTATTTCA 4260
QY 1328 SerValIleAlaSerThrProIleLeuProGlnAlaAlaGlyProThrSerThrProLeu 1347
Db 4261 TCAGTGATGATCTTACTCTCATTTCTTCCCAACAGCAGAGACCTTCTTCAACCTTTA 4320
QY 1348 LeuProGlnValProSerIleProProLeuValGlnProValAlaAsnValProAlaVal 1367
Db 4321 TTACCCCAAGTACTATATTCACCCCTTGTGTACAGCTGTGTGCAATGTGCTGCTGTA 4380
QY 1368 GlnGlnThrLeuIleHisSerGlnProGlnProAlaLeuLeuProAsnGlnProHisThr 1387
Db 4381 CAGCAGACACTAAATTCATAGTCAAGCTCAACAGCTTGTCTTCCCAACAGCCCATACT 4440
QY 1388 HisCysProGlnValAspSerAspThrGlnProValAlaProGlyIleAspAspIleLys 1407
Db 4441 CATTTCTTCAAGTAGTCTGATACCAACCAAACTCTGTGAATTTGATGAATAGAAAG 4500
QY 1408 ThrLeuGlnGluLysLeuArgSerLeuPheSerGluHisSerSerSerGlyAlaGlnHis 1427
```

Db 4501 ACTAGAGAAAAGCTCGGCTCTCTTCAGTGAACACAGCTACATCGAGCTCAGCAT 4560
 Qy 1428 AlaSerValSerLeuGluThrSerLeuValIleGluSerThrValThrProGlyTlePro 1447
 Db 4561 GCCCTGCTCTACCTGAGACCTCACTACATAGAGACCTGTACACCGAGGATCCCA 4620
 Qy 1448 ThrThrAlaValAlaProSerLysLeuLeuThrSerThrThrCysLeuProPro 1467
 Db 4621 ACTACTGCTGTGACCAAGCAAACTCTGACTTCTACCAAGACTTGTCTTACCA 4680
 Qy 1468 ThrAsnLeuProLeuGlyThrValAlaLeuProValThrProValThrProGlyGln 1487
 Db 4681 ACCAATTACACACAGAAAGAGTTGCTTGCAGATTACACAGAGGTGCACACTGGGCAA 4740
 Qy 1488 ValSerThrProValSerThrThrThrSerGlyValLysProGlyThrAlaProSerLys 1507
 Db 4741 GTTCTTACCCAGTACGACTACTACATACAGAGAAACCTGGAACTGCTCTCCAG 4800
 Qy 1508 ProProLeuThrLysAlaProValLeuProValGlyThrGluLeuProAlaGlyThrLeu 1527
 Db 4801 CCACCTCTACTAAGGCTCCGGTGTGCTGACAGTGGGTACTGAACTTCCAGCAGTACTCTA 4860
 Qy 1528 ProSerGluGlnLeuProProPheProGlyProSerLeuThrGlnSerGlnGlnProLeu 1547
 Db 4861 CCCAGCGAGCAGCTGCACCTTTCAGAGACCTTCTTACCCAGTCCAGCAACCTCTA 4920
 Qy 1548 GluAspLeuAspAlaGlnLeuAspArgThrLeuSerProGlu**IleThrValThrSer 1567
 Db 4921 GAGGATCTTGATGCTCAATTGAGAAAGAACTTGTGAGAGATATACAGTGCCTCT 4980
 Qy 1568 AlaValGlyProValSerMetAlaAlaProThrAlaIleThrGluAlaGlyThrGlnPro 1587
 Db 4981 GCGGTGGTCTCTGTCCATGCGGCTCCAAACAGCAATCAGAAAGCAAGCAAGCT 5040
 Qy 1588 GlnLysGlyValSerGlnValLysGlnGlyProValLeuAlaThrSerSerGlyValAlaGly 1607
 Db 5041 CAGAAAGGTGTTTCTCAAGTCAAAAGAGCCCTGCTAGCAACTAGTTCCAGAGTCTGCT 5100
 Qy 1608 ValPheLysMetGlyArgPheGlnValSerValAlaAlaAspGlyValGlnLysGlnLys 1627
 Db 5101 GTTTTAAAGATGGAGCAATTCAGATTCCTGTTGAGAGAGAGGTCGCAAAAGAGGT 5160
 Qy 1628 LysAsnLysSerGluAspAlaLysSerValHisPheGluSerSerThrSerGluSerSer 1647
 Db 5161 AAAAATAAGTCAAGAAAGTCAAGTCTGTCTCATTTGAATCCAGACCTCAGAGTCTCA 5220
 Qy 1648 ValLeuSerSerSerSerProGluSerThrLeuValLysProGluProAsnGlyTleThr 1667
 Db 5221 GTGCTATCAAGTAGTAGTCCAGAGAGTACCTTGTGAAACCAGACCCGAAATGGCATTAAC 5280
 Qy 1668 IleProGlyLysSerSerAspValProGluSerAlaHisLysThrThrAlaSerGlnAla 1687
 Db 5281 ATCCCTGATATCTTTCAGATGTGCCAGAGAGTGCACAAAACTACTGCTCAAGAGCA 5340
 Qy 1688 LysSerAspThrGlnGlnProThrLysValGlyArgPheGlnValThrThrAlaAsn 1707
 Db 5341 AAGTAGACACTGGGCACTCAAGAGTGGACCTTTTCAAGTACCAACTACACAAAC 5400
 Qy 1708 LysValGlyArgPheSerValSerLysThrGluAspLysIleThrAspThrLysLysGln 1727
 Db 5401 AAAGGGGCTGCTTCTCTGATCAAAAACCTGAGCAAGATCACTGACACAAAGAA 5460
 Qy 1728 GlyProValAlaSerProProPheMetAspLeuGlnGlnAlaValLeuProAlaValIle 1747
 Db 5461 GGACCAAGGCAATCTCTCTTAAATGGATTTGGAAACAAGCTGCTCTGCTGTGATA 5520
 Qy 1748 ProLysLysGluLysProGluLeuSerGluProSerHisLeuAsnGlyProSerSerAsp 1767
 Db 5521 CCAAGAAAGAGAAAGCTGAACTGTCAAGCTTCAACTTAATAGGGGCGTCTTGTAC 5580
 Qy 1768 ProGluAlaAlaPheLeuSerArgAspValAspAspGlySerGlySerProHisSerPro 1787

Db 5581 CCGAGGCCCGCTTTTAAAGTAGGATGTGAGATGTTCCGGTAGTCCAGCTCGCCC 5640
 Qy 1788 HisGlnLeuSerSerLysSerLeuProSerGlnAsnLeuSerGlnSerLeuSerAsnSer 1807
 Db 5641 CATCAGCTGAGCTCAAAAGAGCTTCTTACCCAGAAATCTAAGTCAAAAGCTTATGATTA 5700
 Qy 1808 PheAsnSerSerLysMetSerSerAspAsnGluSerAspIleGluAspGluAspLeuLys 1827
 Db 5701 TTTAACCTCTTTCATAGATGAGCAATGATGATGATGATGATGATGATGATGATGATGAT 5760
 Qy 1828 LeuGluLeuArgGluLeuArgAspLysHisLeuLysGlnIleGlnAspLeuGlnSerArg 1847
 Db 5761 TTAGAGCTGCACACACTACGATTAACATCTCAAGAGATTCCAGACCTTCAAGAGTGC 5820
 Qy 1848 GlnLysHisGlnIleGluSerLeuThrLysThrLysLeuGlyLysValProProAlaValIle 1867
 Db 5821 CAGAAAGCTGAAATTTGAATCTTTGTATACCAACTGGGCAAGAGTGCCTCTGTATTT 5880
 Qy 1868 IleProProAlaAlaProLeuSerGlyArgArgArgArgProThrLysSerLysGlySer 1887
 Db 5881 ATTCGCCAGCTGCTCCCTTTCAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5940
 Qy 1888 LysSerSerArgSerSerSerSerLysLysAsnLysSerProGlnLeuSerGlyAsnLeuSer 1907
 Db 5941 AAATCTGATGAAACAGTCTCTTGGGAAATAAAGCCCCAGCTTCAAGTAACTGTCT 6000
 Qy 1908 GlyGlnSerAlaAlaSerValLeuHisProGlnGlnThrLysHisProProGlyAsnIle 1927
 Db 6001 GGTCAAGATGACAGCTTCAAGTCTTGTGACCCCAAGAGAGAGAGAGAGAGAGAGAGAG 6060
 Qy 1928 ProGluSerGlyGlnAsnGlnLeuLeuGlnProLeuLysProSerProSerSerAspAsn 1947
 Db 6061 CCAGAGTCCGGGAGAAATACAGCTGTACAGCCCTTAAAGCCATCTCCCTCAGTACACAC 6120
 Qy 1948 LeuTyrSerAlaPheThrSerAspGlyValIleSerValProSerLeuSerAlaProGly 1967
 Db 6121 CTCTATTCAGCTTCCACAGATGATGTGCCATTTTCAGTACCAAGCCTTTCGCTCCAGT 6180
 Qy 1968 GlnGlyThrSerSerThrAsnThrValGlyValAlaThrValAsnSerGlnAlaAlaGlnAla 1987
 Db 6181 CAAGAGACACAGACCAAAACATGTTGTGGGCAACAGTGAACAGCCAGCCGACCT 6240
 Qy 1988 GlnProProAlaMetThrSerSerArgLysGlyThrPheThrAspAspLeuHisLysLeu 2007
 Db 6241 CAGCTCTGTCATGACGCTCCAGCAGAGAGGACATTCACAGATGACTTGCACAAATGG 6300
 Qy 2008 ValAspAsnThrPalaArgAspAlaMetAsnLeuSerGlyArgArgGlySerLysGlyHis 2027
 Db 6301 GTAGACAATTTGGGCGCGAGATGCCATGATCTTCAAGCAGAGAGAGAGAGAGAGAGAG 6360
 Qy 2028 MetAsnTyrGlnGlyProGlyMetAlaArgLysPheSerAlaProGlyGlnLeuCysIle 2047
 Db 6361 ATGATTTATGAGGCGCTCGAAATGGCAAGAAAGTTCTGTGACCTGGGCAACTGTGATC 6420
 Qy 2048 SerMetThrSerAsnLeuGlyLysSerAlaProIleSerAlaAlaSerAlaThrSerLeu 2067
 Db 6421 TCCATGACCTCGAAGCTGGGTGGCTCTGCCCACTCTGAGAGATCACTGACCTCTCTA 6480
 Qy 2068 GlyHisPheThrLysSerMetCysProProGlnGlnTyrGlyPheProAlaThrProPhe 2087
 Db 6481 GGTCACTTCAACCAAGTCTATGTGCCCCCAAGCAAGATGAGCTTTCAGCTTACCCCATTT 6540
 Qy 2088 GlyAlaGlnTyrPheSerGlyThrGlyGlyProAlaProGlnProLeuGlyGlnPheGlnPro 2107
 Db 6541 GGCGCTCAATGAGAGGAGCGGTGGCCCAAGCAAGAGCAAGCAAGTGGCAAGTCCAACT 6600
 Qy 2108 ValGlyThrAlaSerLeuGlnAsnPheAsnIleSerAsnLeuGlnLysSerLysSerAsn 2127
 Db 6601 GTGGGAAGTGCCTCTCTCAGAAATTTCAACATCAAGCAATTTGAGAAATTCATCAGCAAC 6660
 Qy 2128 ProProGlySerAsnLeuArgThrThr 2136
 Db 6661 CCCCAGGCTCCAACTCGAGCACT 6687

QY 401 GluSerValAspValIleTyrAlaPheGlyMetCysMetLeuGluMetAlaThrSerGluTyr 420
DB 1201 GAATCCGTTGACGTTTATGCTTTGGGATGTGCATGCTTGAGATGGGCTACATCTGAATAT 1260
QY 421 ProTyrSerGluCysGlnAsnAlaAlaGlnIleTyrArgValThrSerGlyValIlys 440
DB 1261 CCTTACTCGGAGTCCAAATGCTCGCAGATCTACCTCGCGACCATGCGGCTGAG 1320
QY 441 ProAlaSerPheAspLysValAlaIleProGluValIleGluGluCysAlaIle 460
DB 1321 CCGCCGATTTTGACAAAGTACATCTGAGAGGAAATTTATGAAAGATGCACTA 1380
QY 461 ArgGlnAsnLysAspGluArgTyrSerIleLysAspLeuLeuAsnHisAlaPhePheGln 480
DB 1381 CGACAAACCAAGATGAAAGATTTCCATCAAAAGCCTTTGAAACCATGCTTCTCCAA 1440
QY 481 GluGluThrGlyValArgValGluLeuAlaGluGluAspAspGlyGluLysIleAlaIle 500
DB 1441 GAGGAAACAGAGATACGGGTGAAATGACAGAGAGATGATGAGAAAAATAGCCATA 1500
QY 501 LysLeuThrLeuArgIleGluAspIleLysLysLeuLysGlyLysTyrLysAspAsnGlu 520
DB 1501 AAATTATGCTACGATTTAGAAATTTAAAGAAATTTAAAGAAATTTAAAGAAATTTAA 1560
QY 521 AlaIleGluPheSerPheAspLeuGluArgAspValProGluAspValAlaGlnGluMet 540
DB 1561 GCTATTGAGTTTCTTTTGAATTTAGAGAGATGCTCCAGAAAGATGTTGCACAAAGAAATG 1620
QY 541 ValGluSerGlyTyrValCysGluGlyAspHisLysThrMetAlaLysAlaIleLysAsp 560
DB 1621 GTACAGCTGGGTATGCTGTGAAGGTGATCACAAGACCATGGCTAAAGCTATCAAGAAC 1680
QY 561 ArgValSerLeuIleLysArgLysArgGluGlnArgGluLeuValArgGluGlnGlu 580
DB 1681 AGAGATCATTAATTAAGAGAAACGAGAGCAGCGAGTTGTGAGGAGAGAGAGAA 1740
QY 581 LysLysLysGlnGluGluSerSerLeuLysGlnGlnValGlnGlnSerSerAlaSerGln 600
DB 1741 AAAAAAAGCAGAGAGAGAGAGAGCTCTCAACAGCAGGTAGAAACATCCAGTCTCCAG 1800
QY 601 ThrGlyIleLysGlnLeuProSerAlaSerThrGlyIleProThrAlaSerThrThrSer 620
DB 1801 ACAGAAATCAAGCAGCTCCCTTCTGCTAGCAGCCGCAATCACTACTCTTACCACTTCA 1860
QY 621 AlaSerValSerThrGlnValGluProGluGluProGluAlaAspGlnHisGlnGluLeu 640
DB 1861 GCTTCAGTTCTACCAAGATGAACCTGAAGAACTGAAGCAATCAATCAATCAACTA 1920
QY 641 GlnTyrGlnGlnProSerIleSerValLeuSerAspGlyThrValAspSerGlyGlnGly 660
DB 1921 CAGTACCGCAACCCAGTATATCTGTATCTGATGGGACGGTTGACAGTGTCAAGGA 1980
QY 661 SerSerValPheThrGluSerArgValSerSerGlnGlnThrValSerTyrGlySerGln 680
DB 1981 TCCCTGCTCTTACAGAAATCTCGAGTGAGCAGCCAAACAGACAGTTTCAATGTTCCAA 2040
QY 681 HisGlnGlnAlaHisSerThrGlyThrValProGlyHisIleProSerThrValGlnAla 700
DB 2041 CATGAACAGGCACATTTACAGGACAGTCCAGGAGCATATACCTTCACTGTCACCA 2100
QY 701 GlnSerGlnProHisGlyValTyrProProSerSerValAlaGlnGlyGlnSerGlnGly 720
DB 2101 CAGTTCAGCCCATGGGGTATATCCACCTCAAGTGTGGCAGGAGGAGCAGGCAAGGCT 2160
QY 721 GlnProSerSerSerSerLeuThrGlyValSerSerSerGlnProIleGlnHisProGln 740
DB 2161 CAGCCATCTCAAGAGCTTAAACAGGGGTTTCATCTCCAAACCATCAACATCTCAG 2220
QY 741 GlnGlnGlnGlyIleGlnGlnThrAlaProProGlnGlnThrValGlnTyrSerLeuSer 760
DB 2221 CAGCAGCAGGAAATACAGACAGCAGCCCTCTCAACAGACAGTGCAGATTCCTACTTCA 2280

QY 761 GlnThrSerThrSerSerGluAlaThrThrAlaGlnProValSerGlnProGlnAlaPro 780
DB 2281 CAGACATCAACCTCCATGAGGAGCCATCTAGCAAGCAAGAGTCAAGCTCAAGCTCCA 2340
QY 781 GlnValLeuProGlnValIleSerAlaGlyLysGln----- 791
DB 2341 CAACTTGGCTTCAAGATACAGCTGGAACAAAGCTTCCAGTTTCCAGCAGTACCAACT 2400
QY 791 ----- 791
DB 2401 ATCCAGGCCAAGCTCAGATTCAGTTCCAGACACAAACCTGGTGTTCAGTCCACTCT 2460
QY 791 ----- 791
DB 2461 GGTCGATATTTCTTCCAGTGGAGACAGCGGCTCCCTACTGCTTGTCTCCAGTACCT 2520
QY 791 ----- 791
DB 2521 GTCTTCAGATTTCCATATCAACTCTCATGTGTCTACGGCTCAGACAGGTTTCTATCC 2580
QY 791 ----- 791
DB 2581 CTTCCTCATCAATGGCAGCTGSCATTACTCAGCTCTGCTCAGCTTGCTTCATCTGCT 2640
QY 791 ----- 791
DB 2641 ACAACAGCTGGATCCCGGGGTATCAACTGTGTCTTACTAGTCACTTCAACCTTCTGT 2700
QY 791 ----- 791
DB 2701 CAGCTGTGACTCAGCTGCCAAGTCAAGTTCACCCAGACTCTTCAACACAGCAGTTGAG 2760
QY 791 ----- 791
DB 2761 TCCATGGGAATACCACTAATCTTGAGACAACTGCTAGGTTCCACTTCTCTGGAGAT 2820
QY 791 ----- 791
DB 2821 GTTCTGTAACAGGCTTCCACCTCGACTGACACACAGTACCAGAGATTTCAAAATAT 2880
QY 791 ----- 791
DB 2881 GCTCCCTTCCAAAGTGGCTTCTGTTCATTCATTCATTCATTCATTCATTCATTCAT 2940
QY 791 ----- 791
DB 2941 CCGACAGAGTCTGGCTACACTGGGTACTTTCACAGCTGTCAGACCTTATGTGAA 3000
QY 791 ----- 791
DB 3001 TCAATCTTTTATGTTCTATGGGTGTGTAGAGAGACAGTTCAGATGTCACACAGCA 3060
QY 792 ----- 792
DB 3061 GGGAGTTTAGCAAGCCCCCATCAATCTCCAGACAGCATTTTGGAGAGTACTCAG 3120
QY 795 GlnValSerGlnValAlaProAlaGluProValAlaValAlaGlnProGlnAlaThrGln 814
DB 3121 GAGTCTCTCAGGTGCTTCTCGCAGAGCCAGTTCAGTACAGACGCCCCAAGCTACCCAG 3180
QY 815 ProThrThrLeuAlaSerSerValAspSerAlaHisSerAspValAlaSerGlyMetSer 834
DB 3181 CCGACCACTTGGCTTCTCTGTAGACAGTCACTTCAGATGATGTGCTTCAGGTATGAGT 3240
QY 835 AspGlyAsnGluAsnValProSerSerSerGlyArgHisGlnGlyTyrArgThrThrLysArg 854
DB 3241 GATGGCAATGAAACGTCCTCATCTTCCAGTGAAGGCTGGAAGAGAACTACAAAAGGG 3300
QY 855 HisTyrArgLysSerValArgSerArgSerArgHisGlnLysTyrThrSerArgProLysLeu 874
DB 3301 CATACCGAAATCTGTAGAGAGTCTCTCGACATGAAATAATTCACGCGCCAAAATTA 3360
QY 875 ArgIleLeuAsnValSerAsnLysGlyAspArgValValGlnCysGlnLeuGluThrHis 894


```
Db 3361 AGAATTGTAAGTTTCAATTAAGAGAGCCGAGTAGTAGATGTCATTAGAGCTCAT 3420
Qy AsnArglySmetValThrPheLysPheAspLeuAspGlyAspAsnProGluGluIleAla 914
Db 3421 AATAGGAAAAATGTTACATTCATAATTTGACCTAGATGGTGACACCCGAGAGATAGCA 3480
Qy ThrIleMetValAsnAspPheIleLeuAlaIleGluArgGluSerPheValAspGln 934
Db 3481 ACAATTATAGTGCAACAATGACTTTATTCTAGCAATAGAGAGAGAGTCTTTGTGATGCA 3540
Qy ValArgGluIleIleGluLysAlaAspGluMetLeuSerGluAspValSerValGluPro 954
Db 3541 GTGCGAGAAATATATTGAAAAAGCTGATGAATGCTCAGTAGAGAGATGTCAGTGTGACCA 3600
Qy GluGlyAspGlnGlyLeuGluSerLeuGlnGlyLysAspAspTyrGlyPheSerGlySer 974
Db 3601 GAGGAGTATCAGAGGAGTTGAGAGTCTACAGAGAAAGATGACTATGGCTTTTCAGGTTCT 3660
Qy GluLysLeuGluGlyGluPheLysGlnProIleProAlaSerSerMetProGlnGlnIle 994
Db 3661 CAGAAATTGGAAGAGAGAGTTCAAAACCAATTCCTGCGTCTTCATGCGCACAGCAATTA 3720
Qy GlyIleProThrSerSerLeuThrGlnValValHisSerAlaGlyArgArgPheIleVal 1014
Db 3721 GGCAATTCCTACCAAGTCTTTAACTCAAGTTGTTCAATTCGCGGAGAGCGGTTTATAGTG 3780
Qy SerProValProGluSerArgLeuArgLeuArgLysValPheProSerGluIleThrAsp 1034
Db 3781 AGTCTGTGCGCAAGAAAGCCGATACGAGAAATCAAAAGTTTCCCGAGGAAATTAACAAT 3840
Qy ThrValAlaAlaSerThrAlaGlnSerProGlyMetAsnLeuSerHisSerAlaSerSer 1054
Db 3841 ACAAGTTGTGCTCTCAACAGCTCAGAGCCCTGGAAGAACTTGTCTCATCTGTCATCAACC 3900
Qy LeuSerLeuGlnGlnAlaPheSerGluLeuArgArgAlaGlnMetThrGluGlyProAsn 1074
Db 3901 CTTAGCTCAACAAGGCTTTTCTGAACCTTAGACGTGCCCAATGACAGAAAGACCCAT 3960
Qy ThrAlaProProAsnPheSerHisThrGlyProThrPheProValValProPheLeu 1094
Db 1095 AOCAGACTCCAAACTTATGATACAGAGCCAAATTTCCAGTAGTACCTCTTCTTA 4020
Qy SerSerIleAlaGlyValProThrThrAlaAlaAlaThrAlaProValProAlaThrSer 1114
Db 4021 AGTAGCATTTGCTGAGTCCCAACACAGACAGCAGACAGACAGTCCCTGCAACAAGC 4080
Qy SerProProAsnAspIleSerThrSerValIleGlnSerGluValThrValProThrGlu 1134
Db 4081 AGCCCTCTATGACATTTCCACATCAGTAATTCAGTCTGAGGTTACAGTGCCCACTGAA 4140
Qy GluGlyIleAlaGlyValAlaThrSerThrGlyValValAlaThrSerGlyLysLeuProIle 1154
Db 4141 GAGGAGAAATGCTGAGGTGGCCACAGCAGAGTGGTGAACCTTCAAGGTGCTCCCAATA 4200
Qy ProProValSerGluSerProValLeuSerSerValValSerSerIleThrIleProAla 1174
Db 4201 CCACCTGTGTGATGATCCACAGTACTTCCAGCGTAGTTTCAAGTATACAAATACCTGCA 4260
Qy ValValSerIleSerThrThrSerProSerLeuGlnValProThrSerThrSerGluIle 1194
Db 4261 GTTCTCTCAATATCTACATCAATCCCGCTCACTCAAGTCCCAATCACAATCAGAGATC 4320
Qy ValValSerSerThrAlaLeuThrProSerValThrValSerAlaAlaThrSerAlaSerAla 1214
Db 4321 GTTGTTCCTATGATACGACGATGATCTTCAGTAACAGTTTCAAGCAACTTCAGCTCTGCA 4380
Qy GlyGlySerThrAlaAlaThrProGlyProLysProProAlaValAlaSerGlnAlaAla 1234
Db 4381 GGGGGCAGTACTGCTACCCAGGTCCTTAGGCTCCAGCTGATAGTATCTCAGACAGGCA 4440
Qy GlySerThrThrValGlyAlaThrLeuThrSerValSerThrThrThrSerPheProSer 1254
Db 4441 GGCAGCACTACTGTGAGAGCCACATTAATCAGTTTCTACACCACTTCATTCACCAAGC 4500
Qy ThrAlaSerGlnLeuSerIleGlnLeuSerSerThrSerThrProThrIleuAlaGlu 1274
Db 4501 ACAGCTTACAGCTGTCCATTCAGCTTACAGCAGTACTTACTTACTTACTTACTTACTTAA 4560
Qy ThrValAlaValSerAlaHisSerLeuAspLysThrSerHisSerSerThrThrGlyLeu 1294
Db 4561 ACCGTGTAGTGGAGCAGCACTACTAGATTAAGATCTCATAGAGTACCACTGATG 4620
Qy AlaPheSerLeuSerAlaProSerSerSerSerSerProGlyAlaGlyValSerSerTyr 1314
Db 4621 GCTTCTCTCCTCTGTGACCAATCTTCCCTTCTCTCTGAGAGCAGAGTGTCTAGTAT 4680
Qy IleSerGlnProGlyGlyLeuHisProLeuValIleProSerValIleAlaSerThrPro 1334
Db 4681 ATTCTCAGCGTGGGCTGATCTCTTGGTCATTCATCAGTATAGCTTCTTCTCT 4740
Qy IleLeuProGlnAlaAlaGlyProThrSerThrProLeuLeuProGlnValProSerIle 1354
Db 4741 ATTCTCCCAAGACAGCAGACCTTCTACACCTTATTAACCCCAAGTACATGATAC 4800
Qy ProProLeuValGlnProValAlaAsnValProAlaValGlnGlnThrLeuIleHisSer 1374
Db 4801 CCACCTTGTATACAGCTGTGCCAATGTCCTGTGTATAGCAGACACTAATTCATAGT 4860
Qy GlnProGlnProAlaAlaLeuLeuProAsnGlnProHisThrHisCysProGluValAspSer 1394
Db 4861 CAGCCTCAACAGCTTGTGCTTCCCAACAGCCCATATCTCATTTGCTCCGAAAGTATCT 4920
Qy AspThrGlnProLysAlaProGlyIleAspAspIleLysThrLeuGluGluLysLeuArg 1414
Db 4921 GATACACACCCCAAGCTCTCGAATTGATACATTAAGACTCTTAGAAGAAAGCTGCGG 4980
Qy SerLeuPheSerGluHisSerSerSerSerGlyAlaGlnHisAlaSerValSerLeuGluThr 1434
Db 4981 TCTCTGTGATGAGTAACAGCTCATCTGAGACTCAGACATGCTGTGTCTGCTGAGAGCC 5040
Qy SerLeuValIleGlnSerThrValThrProGlyIleProThrThrAlaAlaProSer 1454
Db 5041 TCACCTAGCATAGAGAGACAGCTCAGACACAGCAATCCCACTACCTGCTGTGACCAAGC 5100
Qy LysLeuLeuThrSerThrThrSerThrCysLeuProProThrAsnLeuProLeuGlyThr 1474
Db 5101 AAACCTCTGACTTTCACCAAGATCTTGTCTTACACCAACAATTTCCACTAGAACCA 5160
Qy ValAlaLeuProValThrProValValThrProGlyGlnValSerThrProValSerThr 1494
Db 5161 GTTGCTTGGCAGTTACACCAAGTGTGCACACTGGGAAAGTTTCTAACCCAGTCAAGCACT 5220
Qy ThrThrSerGlyValLysProGlyThrAlaProSerLysProProLeuThrLysAlaPro 1514
Db 5221 ACTACATCAGAGAGAACTGGAATGCTCTCCCAAGCCACTTCAACTAAGAGCTTCG 5280
Qy ValLeuProValGlyThrGlnLeuProAlaGlyThrLeuProSerGlnGlnLeuProPro 1534
Db 5281 GTGCTGCAAGGGTACTGAATCTTCCAGCAGTACTTACCCAGCAGAGCTGCCACT 5340
Qy PheProGlyProSerLeuThrGlnSerGlnProLeuGlnProLeuGlnAspAlaGlnLeu 1554
Db 5341 TTTCCAGAGACTTCTCTAACCAGTCCAGCAACTCTAGAGAGATCTTGATGCTCAATTG 5400
Qy ArgArgThrLeuSerProGlu***IleThrValThrSerAlaValAlaGlyProValSerMet 1574
Db 1555 ArgArgThrLeuSerProGlu***IleThrValThrSerAlaValAlaGlyProValSerMet 1574
Db 5401 AGAAGAACACTTATCTCAGAGATGATACATGACTTCTGTGGTGTGCTGTGCTCATG 5460
Qy AlaAlaProThrAlaIleThrGluAlaGlyThrGlnProGlnLysGlyValSerGlnVal 1594
Db 1575 GCGGCTCAACAGCAATCAAGAGCAGAGCAACAGCTCAGAGAGGTGTTTCAAGTC 5520
Qy LysGlnGlyProValLeuAlaAlaThrSerSerGlyAlaGlyValPheLysMetGlyArgPhe 1614
Db 1595 AAAGAGGCCCTGTCTTAGCAACTAGTTCAAGAGACTGGTGTGTTTAAAGATGGAGCATTT 5580
```


QY 1615 GlnValSerValAlaIaAspGlyAlaGlnLysGlnGlyLysAsnLysSerGlnAspAla 1634
 Db 5581 CAGGTTTCTGTGACGACGAGCGGTGCCAAGAAAGGGTAAATAATGATCAGAAAGATCA 5640
 QY 1635 LysSerValHisPheGlnSerSerThiSerGlnSerValLeuSerSerSerPro 1654
 Db 5641 AAGCTGTTCATTTTGATTCAGCACTCAGAGTCCCTCAGTATCAAGTAGAGTCA 5700
 QY 1655 GlnSerThiLeuValLysProGlnProGlnGlyLethrIleProGlnLysSerAsp 1674
 Db 5701 GAGAGTACTTGTGTAACACGAGCCGATGCAATACCTCCGTGATCTCTTCAGAT 5760
 QY 1675 ValProGlnSerAlaHisLysThiThiAlaSerGlnAlaLysSerAspThiGlyLanPro 1694
 Db 5761 GTGCACGAGAGTCCCAACAACTCTGCTCAGAGGCAAACTCAGACACTGGGCAAGCT 5820
 QY 1695 ThrLysValGlyArgPheGlnValThiThiAlaAsnLysValGlyArgPheSerVal 1714
 Db 5821 ACCAAGGTGGAGCTTTCAAGGTGACAACTACAGCAACAAAGTGGGTGCTTCTGTGA 5880
 QY 1715 SerLysThiGlnAspLysIleThiAspThiLysLysGlnGlyProValAlaSerProPro 1734
 Db 5881 TCAAAACAGAGACAGATGATCTGACACAAAGAAAGAGACAGTGGCACTCTCTCTCT 5940
 QY 1735 PheMetAspLeuGlnGlnAlaValLeuProAlaValIleProLysLysGlnLysProGln 1754
 Db 5941 TTTATGATTTGGAACAAAGCTGTTCTCTCTGCTGATACCAAGAAAGAGAGAGCTGAA 6000
 QY 1755 LeuSerGlnProSerHisLeuAsnGlyProSerSerAspProGlnAlaIlePheLeuSer 1774
 Db 6001 CTGTGAGAGCTTCATCATCTAAATGGCGCTCTTCTGACCGGAGGCGCTTTTATTAAGT 6060
 QY 1775 ArgAspValAspAspGlySerGlySerProHisSerProHisGlnLeuSerSerLysSer 1794
 Db 6061 AGGAGTGTGATGATGTTCCGGTAGTCAACCTCGCCCACTAGCTAGCTCAAAAGGC 6120
 QY 1795 LeuProSerGlnAsnLeuSerGlnSerLeuSerAsnSerPheAsnSerSerLysSer 1814
 Db 6121 CTTCCTAGCCAGAACTTAAGTCAAAAGCTTAGTAACTTAATTAATTAATTAATTAATTA 6180
 QY 1815 SerAspAsnGlnSerAspIleGlnAspGlnAspLeuLysLeuGlnLeuArgLeuArg 1834
 Db 6181 AGCGACATGAGTCAAGATATGAGATGAGAGCTTAAGCTTAAGAGCTCGACGACTAGCA 6240
 QY 1835 AspLysHisLeuLysGlnIleGlnAspLeuGlnSerArgGlnLysHisGlnIleGlnSer 1854
 Db 6241 GATTAACATCTCAAGAGATTCAGAGCTGCGAGAGTCCCGCAAGACATGAATGAACTCT 6300
 QY 1855 LeuTyThrLysLeuGlyLysValProProAlaValIleIleProProAlaAlaProLeu 1874
 Db 6301 TTGTATACCAAACTGGGCAAGGTGCCCCCTGTGTATTAATTCCTCCCAAGCTGCTCCCT 6360
 QY 1875 SerGlyArgArgArgArgProThiLysSerLysGlySerLysSerSerArgSerSerSer 1894
 Db 6361 TCAGGGAGAGAGACGACGACCACTAAAGCAAAAGGACCAAAATCTAGTCAAGCAGTCTCC 6420
 QY 1895 LeuGlyAsnLysSerProGlnLeuSerGlyAsnLeuSerGlyGlnSerAlaAlaSerVal 1914
 Db 6421 TTGGGGAAATAAAGCCCGAGCTTCAGGTAACTGCTGTGCAAGTGCAGGCTTCAATC 6480
 QY 1915 LeuHisProGlnGlnThiLeuHisProProGlnAsnIleProGlnSerGlyGlnAsnGln 1934
 Db 6481 TTGCACCCCGACAGACCTCCACCTCTCTGCAACATCCAGAGTCCGGGCGAGATCAG 6540
 QY 1935 LeuLeuGlnProLeuLysProSerProSerSerAspAsnLeuTySerAlaPheThiSer 1954
 Db 6541 CTGTATACAGCCCTTAAGCATCTCTCCAGTGAACAACCTATATCAAGCTTCAACAGT 6600
 QY 1955 AspGlyAlaIleSerValProSerLeuSerAlaProGlnGlnGlyThiSerSerThiSer 1974
 Db 6601 GATGTGTCCATTTCAATCAAGAGCTTTCTGTCCAGGTCAAGGAAACGACGACGACACAC 6660

QY 1975 ThrValGlyAlaThiValIaSerSerGlnAlaIaGlnAlaGlnProProAlaMetThiSer 1994
 Db 6661 ACTGTGGGGCAACAGTGAACAGCCAGCCGCAAGCTCAGCTCTCTGACATGACGTCC 6720
 QY 1995 SerArgLysGlyThiThiPheThiAspLeuHisLysLeuValAspAsnTTPAlaArgAsp 2014
 Db 6721 AGCAGAAAGGACATTCACAGATGACTTGGCAAAATGGTGTGACAAATGGGCGCGAGAT 6780
 QY 2015 AlaMetAsnLeuSerGlyArgArgGlySerLysGlyHisMetAsnTyArgLysProGly 2034
 Db 6781 GCCATGATTTCTCTAGGACGAGAGAGAGAAAGCAAGGCAATGATTAACGAGGGCTTGA 6840
 QY 2035 MetAlaArgLysPheSerAlaProGlyGlnLeuCysIleSerMetThiSerAsnLeuGly 2054
 Db 6841 ATGGCAAGAAAGTCTCTGACACTGGGCACTGTGATCTCATGACCTGCAACTGGCT 6900
 QY 2055 GlySerAlaProIleSerAlaIleSerAlaThiSerLeuGlyHisPheThiLysSerMet 2074
 Db 6901 GGCTCTGCCCCCATCTCTGCGACATCGACTACCTCTTAGGTCACTTCAACCAAGTCTAG 6960
 QY 2075 CysProProGlnGlnTyArgLysPheProAlaThiProPheGlyAlaGlnTTPSerGlyThi 2094
 Db 6961 TGCCCCCACAGCAGTATGGCTTTCCAGCTAACCCCATTTGGCGCTCAATGAGTGGACG 7020
 QY 2095 GlyGlyProAlaProGlnProGlnProLeuGlyGlnPheGlnProValGlyThiAlaSerLeuGln 2114
 Db 7021 GGTGGCCAGCAGCAGCAGCAGCAGCTTGGCCAGTTCCAACTGTGGGAAGTGGCTTGGAG 7080
 QY 2115 AsnPheAsnIleSerAsnLeuGlnLysSerIleSerAsnProProGlySerAsnLeuArg 2134
 Db 7081 AATTTCACAACTACAGCAATTTGCAAGAAATCCATGACAGCAACCCCGAGCTCCAACTGGCG 7140
 QY 2135 ThrThr 2136
 Db 7141 ACCTACT 7146

RESULT 5
 ACH03764
 ID ACH03764 standard; DNA; 7149 BP.
 XX
 AC ACH03764;
 XX
 DT 26-SEP-2003 (first entry)
 XX
 DE DNA encoding novel human protein kinase #1.
 XX
 KM Human; kinase; ds; gene; cosmetic application; nutraceutical application.
 XX
 OS Homo sapiens.
 XX
 PN US6541252-B1.
 XX
 PD 01-APR-2003.
 XX
 PF 14-MAY-2001; 2001US-00854856.
 XX
 PR 19-MAY-2000; 2000US-0206015P.
 XX
 PA (TEXT-) LEXICON GENETICS INC.
 XX
 PI Walke DW, Hildun E, Donoho G, Turner CA;
 XX
 DR MPI; 2003-575927/54.
 DR P-PsDB; ABO44381.
 XX
 PT New nucleic acid encoding novel human proteins, useful in cosmetic and
 XX nutraceutical applications.
 XX
 PS Disclosure, Page; 11pp; English.
 CC The invention relates to a new isolated nucleic acid encoding a novel
 CC human protein kinase. The nucleic acid is useful in cosmetic and
 CC nutraceutical applications. The present sequence represents DNA encoding

Db	1921	CAGACACGACCAACCGATATCTGTATCTATCTATCTGAGGAGCGTTGACAGTGTACAGGGA	1980	
Qy	661	SerSerValPheThrGluSerArgValSerSerGlnGlnThrValSerTyrGlySerGln	680	
Db	1981	TCCTCTGCTTTCACAGAACTCTGAGTGGACGACCAAGACAGTTTCAATATGTTCCCA	2040	
Qy	681	HisGlnGlnAlaHisSerThrGlyThrValProGlyHisIleProSerThrValGlnAla	700	
Db	2041	CATGACAGCGACATTTCTACAGGACAGCTCCAGGGGATATACCTTCTACTGTCCAGCA	2100	
Qy	701	GlnSerGlnProHisGlyValTyrProProSerSerValAlaGlnGlyGlnSerGlnGly	720	
Db	2101	CAGTCTCAGCCCCATGGGTATATCCACCTCAAGTGTGGCAGGGGCAAGGCGAGCGGT	2160	
Qy	721	GlnProSerSerSerSerSerLeuThrGlyValSerSerSerGlnProIleGlnHisProGln	740	
Db	2161	CAGCCATCTCTAAGTATAGCTTACAGGGGTTTCATCTCCAAACCAATCAACATCTTCAG	2220	
Qy	741	GlnGlnGlnGlyIleGlnGlnThrAlaProProGlnGlnThrValGlnTyrSerLeuSer	760	
Db	2221	CAGCAGCAGGGAATACAGCAGACAGCCCTCTCAAGACAGAGTGGCATATTCATCTTCA	2280	
Qy	761	GlnThrSerThrSerSerGluAlaThrThrAlaGlnProValSerGlnProGlnAlaPro	780	
Db	2281	CAGCATCAACCTCCAGTGGGCCACTACTGCACAGCCAGTGTGACGCTCAAGCTTCA	2340	
Qy	781	GlnValLeuProGlnValAlaSerAlaGlyValGln-----	791	
Db	2341	CAGTCTTGCTCTCAAGTATCAGCTGGAAAACAGCTTCCAGTTCCCAAGCCAGTACCACT	2400	
Qy	791	-----	791	
Db	2401	ATCCAGGCGAAGCTCAGATCCAGTTGGCAGACAAACCCGTGGTTGTTCCAGTCCACTCT	2460	
Qy	791	-----	791	
Db	2461	GGTGCATATTTCTTCCAGTGGGACAGCCGCTCCCTACTCCCTGCTCCCTCAGTACCT	2520	
Qy	791	-----	791	
Db	2521	GTCCTCAGATTTCCCAATACATCACTCTCATGTGTCTACGGCTCAGACAGGTTTCTATCC	2580	
Qy	791	-----	791	
Db	2581	CTTCCCATCAAAATGGGACGCTGGATTACTCAGCCTGTCTCAGTTGGCTTCTATCTGCT	2640	
Qy	791	-----	791	
Db	2641	ACAACAGCTGCGATCCCGGGGGTATCAACTGTGTTCTTACTAGCTTCCAACCTTCTG	2700	
Qy	791	-----	791	
Db	2701	CAGCCTTGACTCAGCTGCCAAGTCAAGTTCAACCAAGCTCTTCAACCAAGCAGCTTCAG	2760	
Qy	791	-----	791	
Db	2761	TTCATGGGAATACAGGTAACCTTGGACAGACGTGTAGATTCCACTTCTCTGGAGAT	2820	
Qy	791	-----	791	
Db	2821	GTTCCTGACAGGGCTTCCACTCGACTGCCACCAAGTACCAAGAGATTCAAAATTT	2880	
Qy	791	-----	791	
Db	2881	GCTCCCTTTCGAAGTGGCTTGTGTTGATTCATTTCTACAGTCTTAAACCTCCCATG	2940	
Qy	791	-----	791	
Db	2941	CCGACAGAACTACTGGCTACACCTGGGTACTTTCCACAGTGGTGGAGCCTTATGTGAA	3000	
Qy	791	-----	791	
Db	3001	TCAAATCTTTAGTTCTTATGGGTGTGTAGAGACAGGTTCAAGTCTCCACACAGA	3060	
Qy	792	-----	794	SerThrGln
Db	3061	GGGAGTTTACGACAGCCCCCATCAATCTCTCCACAGACAGGTTTGGAGAGTACTGAG	3120	
Qy	795	GlyValSerGlnValAlaProAlaGluProValAlaValAlaGlnProGlnAlaThrGln	814	
Db	3121	GGAGTCTCTCAGGTTGCTCTCCGACAGACAGTTTCAGTACAGACAGCCCAAGTATCCAG	3180	
Qy	815	ProThrThrLeuAlaSerSerValAspSerAlaHisSerAspValAlaSerGlyMetSer	834	
Db	3181	CCGACCACTTGGCTTCTCTGTAGACAGTGCACATTCAGATGTTGCTTCAGGTATGAGT	3240	
Qy	835	AspGlyAsnGluAsnValProSerSerSerGlyValHisGlyGluValGlyThrThrLeuArg	854	
Db	3241	GATGGCAATGAAAGTCCCATCTTCCAGTGGAAAGGCAATGAAAGAACTACAAACCG	3300	
Qy	855	HisTyrArgLysSerValArgSerArgSerArgHisGlnLysThrSerArgProLysLeu	874	
Db	3301	CATTACCGAAATCTGTAAAGAGTGGCTCTGACATGAAAAAATTACAGCCCAAAATTA	3360	
Qy	875	ArgIleLeuAsnValSerAsnLysGlyAspArgValValGlyCysGlnLeuGlnThrHis	894	
Db	3361	AGATTATTGAATGTTTCAAAATTAAGAGACCGAGTATGAAATGTCAATTATGACATCAT	3420	
Qy	895	AsnArgLysMetValThrPheLysPheAspLeuAspGlyAspAsnProGlnGlnIleAla	914	
Db	3421	AATGGAAATGTTTCAATTCATTAATTTGACCTTAATGTGACACCCCGAGAGATACCA	3480	
Qy	915	ThrIleMetValAsnAsnAspPheIleLeuAlaIleGluArgLysPheValAspGln	934	
Db	3481	ACATTATGTGTAACATGACTTATTCTTGCAATAGAGAGAGAGTGTGTGTGATCA	3540	
Qy	935	ValArgGluIleIleGlnLysAlaAspGluMetLeuSerGluAspValSerValGluPro	954	
Db	3541	GTCGAGAAATTTATGAAAAAGCTGATGAATGTCTGAGTGTGAGTGTGAGTGTGAAACA	3600	
Qy	955	GlnGlyAspGlnGlyLeuGlnSerLeuGlnGlyLysAspAspTyrGlyPheSerGlySer	974	
Db	3601	GAGGTATACAGGATTTGAGAGTCTTACAAAGAAAGATGACTATGGCTTTTACAGTCT	3660	
Qy	975	GlnLysLeuGlnGlyGluPheLysGlnProIleProAlaSerSerMetProGlnGlnIle	994	
Db	3661	CAGAAATTTGGAGAGAGTTCAACCAACCAATTCCTCGCTTCCATGCCACACAAATA	3720	
Qy	995	GlyIleProThrSerSerLeuThrGlnValAlaHisSerAlaGlyArgPheIleVal	1014	
Db	3721	GGCATTTCTTACCACTTTTAACTCAAGTTGTTCTCTGGGGAAGCGGTTTATATGTG	3780	
Qy	1015	SerProValProGlnSerArgLeuArgLysSerLysValPheProSerGluIleThrAsp	1034	
Db	3781	AGTCTTGGCAGAAACCCGATTTGAGAAATCAAAAGTTTCCCAAGTAAATATACGAT	3840	
Qy	1035	ThrValAlaAlaSerThrAlaGlnSerProGlyMetAsnLeuSerHisSerAlaSerSer	1054	
Db	3841	ACAGTCTGCTGCTTACAGCTCAGAGCCCTGGAATGAACTTGTCTCACTGCACTATCC	3900	
Qy	1055	LeuSerLeuGlnGlnAlaPheSerGluLeuArgArgAlaGlnMetThrGlnGlyProAsn	1074	
Db	3901	CTTAGTCTACAAAGGCTTTTTCGAACCTTAAAGTGTCCCAAAAGACAGAAAGGCCAAAY	3960	
Qy	1075	ThrAlaProProAsnPheSerHisThrGlyProThrPheProValAlaProProPheLeu	1094	
Db	3961	ACAGCACCCTCCAACTTATGCTATACAGACCAACATTTCCAGTATGATCTCTTTCTTA	4020	
Qy	1095	SerSerIleAlaGlyValProThrThrAlaAlaThrAlaProValProAlaThrSer	1114	
Db	4021	AGTAGCATTTGCTGAGTCCCAACACAGCAGCAACCAAGACCAAGTCCCTGCAACAGC	4080	
Qy	1115	SerProProAsnAspIleSerThrSerValIleGlnSerGluValThrValProThrGlu	1134	
Db	4081	AGCCCTCTATATGACATTTCCATCATCATGATTAATTCAGTATGAGTTTACAGTCCCACTGAA	4140	

[illegible]

QY	1495	ThrhSergValLysProGlyThrlaProSeriLysProPoluThrlValaPro	1514
Db	5221	ACTCAATCAGAGTGAATCTTGAATCTGCTCTCCACAGCCACTTCACTAGAGCTTCG	5280
QY	1515	ValLeuProValGlyThrGluLeuProAlaGlyThrLeuProSergGluLeuPro	1534
Db	5281	GTGCTGCAGGGGTACTGAACCTTCCAGAGGTACTCAACCACCGAGCACTCCACCT	5340
QY	1535	PheProGlyProSeriLeuThrGlnSeriGlnProLeuGluAspLeuAspAlaGlnLeu	1554
Db	5341	TTTCCAGACCTTCTCAACCCAGTCCAGAACCTCTAGAGGATCTTGATGCTCAATG	5400
QY	1555	ArgArgThrLeuSeriProGlu**IleThrValThrSeriAlaValGlyProValSeriMet	1574
Db	5401	AGAGAACAACCTTAGTCCAGAGATATATCAAGATGACTTCTGCGTTGGTCCGTGTCATG	5460
QY	1575	AlaAlaProThrAlaIleThrGluAlaGlyThrGlnProGlnLysGlyValSeriGlnVal	1594
Db	5461	GCGGCTCCAAACAGAAATCAACAGAACAGGAAACAGGCTCAGAAAGGGGTTCCTCAAGTC	5520
QY	1595	LysGluGlyProValIleuAlaThrSeriSeriGlyAlaGlyValaPheLysMetGlyArgPhe	1614
Db	5521	AAAGAAAGCCCTGCTCCTGCAACTAGTTCAAGAGCTGGTGTTTTAAGATGGACGATTT	5580
QY	1615	GlnValSeriValaAlaAlaAspGlyValaGlnLysGluGlyLysAsnLysSeriGluAspAla	1634
Db	5581	CAGCTTCTCTGTCAGACGACGCGCCACAGAAAGAGGTAAATTAAGTCAAGAGATGCA	5640
QY	1635	LysSeriValHisPheGluSeriSeriThrSeriGluSeriValaLeuSeriSeriSeriPro	1654
Db	5641	AAGTCTGTTCAATTTGATCCAGCACTCAAGTCTCAAGTCTCAAGTCAAGTCAAGTCA	5700
QY	1655	GluSeriThrLeuValLysProGluProAsnGlyIleThrIleProGlyIleSeriSeriAsp	1674
Db	5701	GAGAGTACCTTGGTGAACACAGACCGAATGGCATTAACCATCCTGGATCTCTTCAGAT	5760
QY	1675	ValIProGluSeriAlaHisLysThrThrlaSeriGluAlaLysSeriAspThrGlnPro	1694
Db	5761	GTGCAAGAGATGGCCCAAAAACTAACGCTCAAGGCAAAAGTCAACATCTGGGACACT	5820
QY	1695	ThrlSeriValaGlyArgPheGlnValaThrThrThrlaAsnLysValaGlyArgPheSeriVal	1714
Db	5821	ACCAAGGTGGACCTTTTCAGGTCAACAACACACAAACAAAGGGGCTCTTCTCTGTA	5880
QY	1715	SeriLysThrGluAspLysIleThrAspThrlLysLysGluGlyProValaIleSeriProPro	1734
Db	5881	TCAAAAACCTGAGCAACAGATCACTCAACAAAGAAAGAGACAGTGGCATCTCTCCT	5940
QY	1735	PheIleSeriLeuGluGlnAlaValLeuProAlaAlaIleProLysLysGluLysProGlu	1754
Db	5941	TTTATGATTTGGAAACAAGCTGTTCTTCTGCTGTATACCAAGAAAGAAACCTCGAA	6000
QY	1755	LeuSeriGluProSeriHisLeuAsnGlyProSeriSeriAspProGluAlaIlePheLeuSeri	1774
Db	6001	CTGTCAAGGCTTCAACATCAATAATGGGCGGCTTCTGACCGGAGGCGCGCTTTTAAAT	6060
QY	1775	ArgAspValaAspAspGlySeriGlySeriProHisSeriProHisGlnLeuSeriLysSeri	1794
Db	6061	AGGATATPGATGTATGGTTCCTCGATAGTCCACATCGGCCCATCAAGCTCAAAAGC	6120
QY	1795	LeuProSeriGlnaLeuSeriGlnSeriLeuSeriAsnSeriPheAsnSeriLysMetSeri	1814
Db	6121	CTTCTTGAAGCAATCTAAGTCAAAAGCTTAGTAATTCATTTAACTCTCTTCAATGAGT	6180
QY	1815	SeriAspAsnGluSeriAspIleGluAspGluAspLeuLysLeuGluLeuArgArgLeuArg	1834
Db	6181	AGCACAATGATGATCAATATCAAGATGAAGAACTTAAGTTAGAGCTGCACACACTACGA	6240
QY	1835	AspLysHisLeuLysGluIleGlnAspLeuGlnSeriArgGlnLysHisGluIleGluSeri	1854
Db	6241	GATTAACAATCTCAAAAGATTCACAGACCTGCAAGGTGCCCAAGACATGAAATTTGAATCT	6300
QY	1855	LeuLysThrlLysLeuGlyLysValProProAlaValIleIleProProAlaIleProLeu	1874

```

Db      6301 TTGTATACCAACTGGGCAAGTGGCCCTGCTGTTATTATTCCTCCAGCTGCTCCCTT 6360
Qy      1875 SerGlyAArgArgArgProThrIysSerIysGlySerIysSerSerArgSerSer 1894
Db      6361 TCAGGGAAGAGACGACGACCACTTAAAGCAAGGACAGCAATCTAGTCGAGCAGTTC 6420
Qy      1895 LeuGlyAenIysSerProGlnLeuSerGlyAenLeuSerGlyGlnSerAlaSerVal 1914
Db      6421 TTGGGGAATTAAGGCCCCAGCTTCAGGTAACTGCTGGTCAGAGTCGAGCTTCAATC 6480
Qy      1915 LeuHisProGlnGlnThrLeuHisProProGlyAenIleProGlySerGlyGlnSerGln 1934
Db      6481 TTGCACCCCGACAGACACCTCCCTCTGGCAATCCAGAGTCCGGGACAGATCAG 6540
Qy      1935 LeuLeuGlnProLeuIysProSerProSerSerAspAsnLeuIysSerAlaPheHisSer 1954
Db      6541 CTGTTACGCCCCCTTAAGCCATCTCTCCAGTGAACAACCTTATTACGCTTCACCGGT 6600
Qy      1955 AspGlyAlaIleSerValProSerLeuSerAlaProGlyGlnGlyThrSerSerThrAsn 1974
Db      6601 GATGGTGCATTTAGTACGACCAAGCTTCTGCTCCAGTCAAGGAAACGACAGCAAAAC 6660
Qy      1975 ThrValGlyAlaThrValAsnSerGlnAlaAlaGlnIleGlnProProAlaMetThrSer 1994
Db      6661 ACTGTTGGGCAACAGTGAACAGCCACGCCCAAGCTCAGCCCTCCCATGACGTCC 6720
Qy      1995 SerArgIysGlyThrPheThrAspAspLeuHisIysLeuValAspAsnThrAlaArgAsp 2014
Db      6721 AGCGAAGAAAGGCACTTCAAGATGACTTGCAACAAGTTGTAGAACATTTGAGCCGAGAT 6780
Qy      2015 AlaMetAsnLeuSerGlyArgArgGlySerIysGlyHisMetAsnTyrGlyIleProGly 2034
Db      6781 GCCATGATCTCTCAGGACGAGAGAGAAAGCAAGGACATGATTTAGAGGCCCTTGA 6840
Qy      2035 MetAlaArgIysPheSerAlaProGlyGlnLeuCysIleSerMetThrSerAsnLeuGly 2054
Db      6841 ATGCAAGAGAAAGTCTCTGCACTGGGCACTGTCATCTCCATGACCTCGAACCTGGGT 6900
Qy      2055 GlySerAlaProIleSerAlaAlaSerAlaThrSerLeuGlyHisPheThrIysSerMet 2074
Db      6901 GGCTCTGCCCCCATCTCTGACGACATCAAGTCACTCTTCAAGTCACTTCAAGGCTATG 6960
Qy      2075 CysProProGlnGlnIysGlyPheProAlaThrProPheGlyAlaGlnIleProSerGlyThr 2094
Db      6961 TGCCCCCGACAGCATGATGGCTTTCAGCTACCCATTTGGCGCTCATGAGTGGAGCG 7020
Qy      2095 GlyGlyProAlaProGlnProLeuGlyGlnPheGlnProValGlyThrAlaSerLeuGln 2114
Db      7021 GGTCGCCGACGACACACACCACTTGGCCAGTTCNAACCTGTGGGAACCTGCTCTTGCAG 7080
Qy      2115 AsnPheAsnIleSerAsnLeuGlnIysSerIleSerAsnProProGlySerAsnLeuArg 2134
Db      7081 AATTTCACATCAGCAATTTCGAAATTCATCAGCAACCCCGACGCTTCACACTGGG 7140
Qy      2135 ThrThr 2136
Db      7141 ACCACT 7146

```

RESULT 6
ACH03778
ID ACH03778 standard; DNA; 6327 BP.

ACH03778;

26-SEP-2003 (first entry)

DNA encoding novel human protein kinase #15.

Human; kinase; ds; gene; cosmetic application; nutraceutical application.

XX Homo sapiens.

```

PN      US6541252-B1.
XX      01-APR-2003.
PD      14-MAY-2001; 2001US-00854856.
XX      19-MAY-2000; 2000US-0206015P.
PR      (LEXI-) LEXICON GENETICS INC.
XX      Walke DW, Hilbun E, Donoho G, Turner CA;
XX      MPI; 2003-575927/54.
DR      P-PSDB; ABO44395.
XX      New nucleic acid encoding novel human proteins, useful in cosmetic and
PT      nutraceutical applications.
XX      Disclosure; Page; 11pp; English.
CC      The invention relates to a new isolated nucleic acid encoding a novel
CC      human protein kinase. The nucleic acid is useful in cosmetic and
CC      nutraceutical applications. The present sequence represents DNA encoding
CC      a novel human protein kinase. Note: The sequence data for this patent did
CC      not form part of the printed specification but was obtained in electronic
CC      format directly from USPTO at
CC      seqdata.uspto.gov/sequence.html?docID=6541252B1
XX      SQ      Sequence 6327 BP; 1757 A; 1697 C; 1451 G; 1420 T; 0 U; 2 Other;

Alignment Scores:
Pred. No.:      0      Length:      6327
Score:      10646.00      Matches:      2108
Percent Similarity:      98.69%      Conservative:      0
Best Local Similarity:      98.69%      Mismatches:      0
Query Match:      98.46%      Indels:      28
DB:      Gaps:      1

US-10-010-720-14 (1-2136) x ACH03778 (1-6327)
Qy      1 MetSerGlyGlyAlaAlaGlnIysSerSerThrProGlySerLeuPheLeuSerPro 20
Db      1 ATGTCGTGGCGGCGCGGACAGAAAGCAGACGACGACTCCCGGTCTCTGTTCTCGCGCG 60
Qy      21 ProAlaProAlaProIysAsnGlySerSerSerSerSerSerValGlyGlnIysLeuGly 40
Db      61 CGGCTCTGCTCCCGCAAGAAAGTCTCCAGCTCCGATTCCTCGTGGGGAAGAACTGGGA 120
Qy      41 AlaAlaAlaAlaAspAlaValThrGlyArgThrGlnIleTyrArgArgArgArgHisThr 60
Db      121 GCGCGGCGCGCGGACGCTGTACCGGACGACGAGAGTACAGGCGCGCGGACACACT 180
Qy      61 MetAspIysAspSerArgGlyAlaAlaAlaThrThrThrThrGlnHisArgPhePhe 80
Db      181 ATGACAAAGACACACCGTGGGCGCGCGGACCACTACCACTGACGACCGCTTCTTC 240
Qy      81 ArgArgSerValIleCysAspSerAsnAlaThrAlaLeuGlnIleuProGlyLeuProLeu 100
Db      241 CGCCGAGCGTCACTCTGGACTCCAAATGCACCTCACTGAGCTTCCCGGCTTCTCTT 300
Qy      101 SerLeuProGlnProSerIleProAlaAlaValProGlnSerAlaProProGluProHis 120
Db      301 TCCTGCCCCGACCGCAGATCCCGCGGCTGTCCGACAGATGCTCCACCGGACCCGAC 360
Qy      121 ArgGlnIleThrValIleThrAlaThrAlaThrSerGlnValAlaGlnIleProProAlaAla 140
Db      361 CGGGAAGAGACCGTGACCGCCACCGCACTTCCAGAGTACCGCCAGCAGCCCTCCAGCCGCT 420
Qy      141 AlaAlaProGlyGlyGlnIleAlaValAlaGlyProAlaProSerThrValProSerSerThr 160
Db      421 GCGGCCCCCTGGGGAACAGGCGGTGCGGCGCTTCCCGCTGACTGTGCCAGAGTACC 480
Qy      161 SerIysAspArgProValSerGlnProSerLeuValGlySerIysGlnIleProProPro 180

```

Db	481	AGCAAGACCCGCCAGTCTCCAGCTTACGCTTGTGGGACCAAGAGACCCGCCG	540
Qy	181	AlaArgSerGlySerGlyGlyGlySerAlaLysGluProGlnGluArgSerGln	200
Db	541	GCGAGAAATGGGAGCGCGCGCGGCGAGCGCCCAAGAGCCACAGAGGAAACGAGCCAGAG	600
Qy	201	GlnSerAspLleGluGluLysGluThrLysAlaValGlyMetSerAsnAspGlyArgPhe	220
Db	601	CAGGATGATATCGAAGACCTGGAGCCCAAGGCCGCGGGGAAAGTGTAAACGATGGCCGCTTT	660
Qy	221	LeuLysPheAspLleGluLleGlyArgGlySerPheLysThrValTyrLysGlyLeuAsp	240
Db	661	CTCAAGTTGACATCGAATCGGACAGAGGCTCTTTAAAGACGGTCTCAAAAGGCTGGAC	720
Qy	241	ThrGluThrThrValGluValAlaATrPCysGluLeuGlnAspArgLysLeuThrLysSer	260
Db	721	ACTGAAACCAACCGTGGAGATCGCGCTGGTGTAACTGCAGGATTCGAAATTAACAAAGCT	780
Qy	261	GluArgGlnArgPheLysGluGluLysGluMetLeuLysGlyLeuGlnHisProAsnLle	280
Db	781	GAGAGGCGAGATTTTAAAGAAAGCTGAAATGTTTAAAGGCTCTTCAGATCCCAATATT	840
Qy	281	ValArgPheTyrAspSerTrpGluSerThrValLysGlyLysLysCysLleValLeuVal	300
Db	841	GTTGATTTTATGATTCCTGGGAAATCCACAGTAAAGAAAGAAAGATGATGTTTGCTG	900
Qy	301	ThrGluLeuMetThrSerGlyThrLeuLysThrTyrLeuLysArgPheLysValMetLys	320
Db	901	ACTGAACTTAGACGCTCTGGAACACTTAAACGATTCGAAAGGTTTAAAGTATGATGAG	960
Qy	321	LleLysValLeuArgSerTrpCysArgGlnLleLeuLysGlyLeuGlnPheLeuHisThr	340
Db	961	ATCAAAAGTTCTAAGAAAGCTGGTGGCGTCAGATCCTTAAAGGCTCTTCATGTTCTTCA	1020
Qy	341	ArgThrProProlLleLleHisArgAspLeuLysCysAspAsnLlePheLleThrGlyPro	360
Db	1021	CGAACTCCACTATCATTCACCGGCACTTAAAGTGCACATCTTATTCACCGGCCCT	1080
Qy	361	ThrGlySerValLysLleGlyAspLeuGlyLeuAlaThrLeuLysArgAlaSerPheAla	380
Db	1081	ACTGGCTCAGTCAGATTTGGAGACCTCGGTGTGGCAACCTCGAAGCGGCTCTTTTGCC	1140
Qy	381	LysSerValLleGlyThrProGluPheMetAlaProGluMetTyrGluGluLysTyrAsp	400
Db	1141	AAGAGTGTGATAGTACCCCGAGAGTTCATGCGCCCTGAGATGTATGAGGAATATGAT	1200
Qy	401	GluSerValAspValTyrAlaPheGlyMetCysMetLeuGluMetAlaThrSerGluTyr	420
Db	1201	GAATCCGTTGACGTTTATGCTTTGGGATGTCATGCTTGAGTGGCTACATTCGAATAT	1260
Qy	421	ProTyrSerGluCysGlnAsnAlaAlaGlnLleTyrArgArgValThrSerGlyValLys	440
Db	1261	CCTTACTCGAGTGGCAAAATGCTGCGACAGATCTACCGTCCGTCGACCAAGTGGGTTGAG	1320
Qy	441	ProAlaSerPheAspLysValAlaLleProGluValLysGluLleGluGlyCysIle	460
Db	1321	CCAGCCAGTTTGAACAAAGTACATTCCTGAGAGGAAGAAATTAATGAAGATGACATA	1380
Qy	461	ArgGluAsnLysAspGluArgTyrSerLleLysAspLeuLeuAsnHisAlaPhePheGln	480
Db	1381	CGACAAAACAAAGATGAAAGATATTCATCAAAAGCCTTTGAAACATGCTTCTTCCAA	1440
Qy	481	GluGluThrTrpGlyValArgValGluLeuAlaGluGluAspAspGlyGlyLysLleAlaIle	500
Db	1441	GAGGAAACAGAGATTCGGGTGGAATTCAGACAGAAAGATGATGAGGAAAAATATGCCATA	1500
Qy	501	LysLeuThrPheLeuArgLleGluAspLleLysLysLeuLysGlyLysTyrLysAspAsnGlu	520
Db	1501	AAATTAATGCGTACGATTTGAGAGATTTATAGAAATTTAAACGCAAAATTCAAAGATTAAGAA	1560
Qy	521	AlaLleGluPheSerPheAspLeuArgAspValProGluAspValAlaGlnGluMet	540

Db	1561	GCATATGACTTTCTTTTGATTAGAGAGAATGCTCCAGAAAGATTGTTCACAAAGAAATG	1620
QY	541	ValGInuSerGlyTYrValCysGluGlyAspHisysThrMetAlaLysAlaIleLysAsp	560
Db	1621	GTAGAGCTCGGGTATGCTGTGTGAAGGTGATTCACAAAGCAATGGCTAAAGCATCAAGAC	1680
QY	561	ArgValSerLeuIleLysArgLysArgGluArgGluArgGluLeuValArgGluGluGlu	580
Db	1681	AGAGATCATTAATTAAAGAGAAACGAGACAGCGCGCGCTGGTGTGATCGGAGACCAAGAA	1740
QY	581	LysLysLysGlnGluGluSerSerLeuLysGlnGluValGluGluSerSerLysGln	600
Db	1741	AAAAAAGCCAGAGAGAGAGAGATGCTCAAAAGAGAGGTAAACATCAAGTCTCCAG	1800
QY	601	ThrGlyIleLysGlnLeuProSerLysSerThrGlyIleProThrAlaSerThrThrSer	620
Db	1801	ACAGGAATCAAGCAGCTCCCTTCTGTCTAGCACCGGCATACCTACTGCTTACCACTTCA	1860
QY	621	AlaSerValSerThrGlnValGluProGluGluProGluAlaAspGlnHisGlnGlnLeu	640
Db	1861	GCTTCAGTCTTCACACAAAGTGAACCTCGAAMACCTGAGGACAGATCAACATCAACAATA	1920
QY	641	GlnTYrGlnGlnProSerLysSerValLeuSerAspGlyThrValAspSerGlyGlnGly	660
Db	1921	CAGTACCAACCAACCGATATCTGTGTATTCTATGTGAAGGATTGACAGTGGTCAAGGA	1980
QY	661	SerSerValPheThrGluSerArgValSerSerGlnGlnThrValSerTYrGlySerGln	680
Db	1981	TCCTCTGTCTTCACAGAACTCGAGTAGACGCCAACAGACAGTTCTATGTGTCCCA	2040
QY	681	HisGluGlnAlaHisSerThrGlyThrValProGlyHisIleProSerThrValGlnAla	700
Db	2041	CATGAACAGGACATCTTCTACAGGACAGTCCAGGAGATATACCTTCTACTGTCCAGCA	2100
QY	701	GlnSerGlnProHisGlyValTYrProProSerSerValAlaGlnGlyGlnSerGlnGly	720
Db	2101	CAGTCTAGCCCCAGGGGATATATCCACCTCCAAGTGG-----	2139
QY	721	GlnProSerSerSerSerLeuThrGlyValSerSerSerGlnProIleGlnHisProGln	740
Db	2139	-----	2139
QY	741	GlnGlnGlnGlyIleGlnGlnThrAlaProProGlnGlnThrValGlnTYrSerLeuSer	760
Db	2140	---CAGCAGGGAATACAGCAGACAGCCCTCTCAACAGACAGTACGATTCACCTTCA	2196
QY	761	GlnThrSerThrSerSerGlnAlaThrThrAlaGlnProValSerGlnProGlnAlaPro	780
Db	2197	CAGACATCAACCTCCAGTGAAGGCCACATGTGCACAGCCAGAGATCGAGCTCAAGCTTCA	2256
QY	781	GlnValLeuProGlnValSerAlaGlyLysGlnSerThrGlnGlyValSerGlnValAla	800
Db	2257	CAGGCTTGCTCAAGTACATACGCTGGGAAACAGAGTACTCAGGGAAGTCTCAGAGTTGCT	2316
QY	801	ProAlaGluProValAlaValAlaGlnProGlnAlaThrGlnProThrThrLeuAlaSer	820
Db	2317	CCTGCAGAGCCAGTTGCAGTGAAGCACACCCCAACCTACCAGCGGACCACTTGGCTTCC	2376
QY	821	SerValIaAspSerAlaHisSerAspValAlaSerGlyMetSerAspGlyValGlnValAspVal	840
Db	2377	TCTGTAGACATGCACATTCAGATGTGCTTCAGAGTATGATGTATGGCAATGAACAGCTC	2436
QY	841	ProSerSerSerGlyArgHisGlnGlyValArgThrThrLysArgHisTYrArgLysSerVal	860
Db	2437	CCATCTTCCAGTGAAGGACATGAAGGAAGAACTACAAACCGCATTCACCAAAATCTGTGA	2496
QY	861	ArgSerArgSerArgHisGlnLysThrSerArgProLysLeuArgIleLeuAsnValSer	880
Db	2497	AGGAGTGCCTCTCCACATGAAGAAAAATCTCAGCGCCAAATTAAGAAATTGTGAATGTTC	2556
QY	881	AsnLysGlyAspArgValValGluCysGlnLeuGluThrHisAsnArgLysMetValThr	900
Db	2557	AATTAAGAGACCGAGTATGTAATGTCAATTAAGTCACTCAATATAGGAATATGTGTACA	2616

QY 901 PheLysPheAspLeuAspGlyAspAsnProGluGluIleAlaThrIleMetValAspAsn 920
Db 2617 TTCAATATTGACCTAGATGATGTGACAAACCCGAGGAGATGACAACTATTATGTGTAACAAT 2676
QY 921 AspPheIleLeuAlaIleGluArgGluSerPheValAspGluValArgGluIleIleGlu 940
Db 2677 GACTTATTCTAGCAATAGAGAGAGAGCTGTTTGATCAAGTCGAGAAATATTATGAA 2736
QY 941 LysAlaAspGluMetLeuSerGluAspValSerValGluProGluGluAspGluGluLeu 960
Db 2737 AAAGTGTAGTAAATGCTAGTAGAGATGTCAGTGTGGAACCAAGAGGATGATCAGGAGATTG 2796
QY 961 GluSerLeuGluGluLysAspAspTyrGlyPheSerGlySerGluLysLeuGluGluGlu 980
Db 2797 GAGAGTCTACAGAGAAAGAGATGACTATGGCTTTTCAGGTTCTCGAAATTTGGAAAGAGAG 2856
QY 981 PheLysGluProIleProAlaSerSerMetProGluGluIleGlyIleProThrSerSer 1000
Db 2857 TTCAAAACCAACCAATTCCTGCTCTTCATGCCACAGCAAAATAGGCATTCCTACCAAGTTCT 2916
QY 1001 LeuThrGluValIleHisSerAlaGlyArgArgPheIleValSerProValProGluSer 1020
Db 2917 TTAACTCAAGTTGTTCATTCTGCGGAAAGCGGTTTATAGTAGCTCTGTCCAGAAAGC 2976
QY 1021 ArgLeuArgGluSerLysValPheProSerGluIleThrAspThrValAlaAlaSerThr 1040
Db 2977 CGATTAGAGATCAAAAGTTTCCCGCAGTGAATTAACAGATACAGTTGCTGCTCTACCA 3036
QY 1041 AlaGluSerProGlyMetAsnLeuSerHisSerAlaSerSerLeuSerLeuGluAla 1060
Db 3037 GCTCAGAGCCCTGGATGATACCTGTCTCCTCATCTGCATCATCTTACCTTACCAACAGGCC 3096
QY 1061 PheSerGluLeuArgArgAlaGluMetThrGluGluProAsnThrAlaProProAsnPhe 1080
Db 3097 TTTCTGAACTTAGACGTCGCCAAATGACAGAGAGCCAAATACAGCACTTCACAACTT 3156
QY 1081 SerHisThrGlyProThrPheProValValProProPheLeuSerSerIleAlaGlyVal 1100
Db 3157 AGTCATACAGAGCAACCACTTCAGTAGTACTCTCTTTCTTAAGTAGATTCGTGAGATC 3216
QY 1101 ProThrThrAlaAlaIleThrAlaProValProAlaThrSerSerProProAsnAspIle 1120
Db 3217 CCAACCAAGCAGCAGCAGCAGCAGCACTCCCTGCAACACAGAGCCCTCTAAATGACACT 3276
QY 1121 SerHisSerValIleGluSerGluValThrValProThrGluGluGlyIleAlaGlyVal 1140
Db 3277 TTCACATCAATTAATTCAGTCTGAGGTTCAGGTGCCACTGAAGAGGAGATTGCTGAGATT 3336
QY 1141 AlaThrSerThrGlyValValThrSerGlyLeuProIleProProValSerGluSer 1160
Db 3337 GCCACAGCAGCAGGCTGAGTCACTTCAGTGTCTCCCATACCACTGTGTCTGAAACA 3396
QY 1161 ProValLeuSerSerValValSerSerIleThrIleProAlaValValSerIleSerThr 1180
Db 3397 CCAGACTCTTCAGCGGTGTTTCAAGTATCAAAATACCTGCAAGTGTCTCAATATCTACT 3456
QY 1181 ThrSerProSerLeuGluValProThrSerThrSerGluIleValValSerSerThrAla 1200
Db 3457 ACATCTCCGCTACTCAAGTCCCAACATCACTGAGATCGTGTCTTCAAGTACACA 3516
QY 1201 LeuThrProSerValThrValSerAlaThrSerAlaSerAlaGlySerThrAlaThr 1220
Db 3517 CTGATACCTTCAGTAAAGTTTCAGCACTTCAGCTTCGCAAGGAGGAGTACTGCTACC 3576
QY 1221 ProGlyProLysProProAlaValValSerGluGluAlaIleGlySerThrValGly 1240
Db 3577 CCAGGTCTTAAGCCTCCAGCTGTAGTATCTCAGCAGCAGCAGCAGCAGCACTACTGTGAGA 3636
QY 1241 AlaThrLeuThrSerValSerThrThrThrSerPheProSerThrAlaSerGluLeuSer 1260
Db 3637 GCCACATTAACTACGTTTCTTACACACACTTCATTCACAGACAGGCTTCACAGCTGTCC 3696

QY 1261 IleGluLeuSerSerSerThrSerThrProThrLeuAlaGluThrValValSerAla 1280
Db 3697 ATTCAGCTTAGACAGCACTACTTCTACTCTTACTTACTGAAACCGTGTAGTACGCA 3756
QY 1281 HisSerLeuAspLysThrSerHisSerSerThrThrGlyLeuAlaPheSerLeuSerAla 1300
Db 3757 CACTCACTTAATAGACATCTCATAGCAGTAACTGATGGCTTCTCCTCTCTGCA 3816
QY 1301 ProSerSerSerSerSerProGlyAlaGlyValSerSerTyrIleSerGluProGlyVal 1320
Db 3817 CCATCTTCTCTTCTCTCTCTGAGCAGAGTGTCTGATTATTTCTTCTGACCTGTGTGG 3876
QY 1321 LeuHisProLeuValIleProSerValIleAlaSerThrProIleLeuProGluAlaAla 1340
Db 3877 CTGCATCTTGGTCATTCATCATGATAGCTTCTACTCTTATTTCTTCCCAAGCAGCA 3936
QY 1341 GlyProThrSerThrProLeuLeuProGluValProSerIleProProLeuValGluPro 1360
Db 3937 GGACCTACTCTTACACCTTATTAACCCAGTACTGTGTATCCCACTTGTGTGACAGCT 3996
QY 1361 ValAlaAsnValProAlaValGluGluThrLeuIleHisSerGluProGluProAlaLeu 1380
Db 3997 GTTCCAAATGCTGCTGCTGTACAGCAGACACTTAATTCATAGTCAAGCTTCACAGCTTTG 4056
QY 1381 LeuProAsnGluProHisSerHisSerProGluValAspSerAspThrGluProLysAla 1400
Db 4057 CTTCCCAACGAGCCCACTACTTGTCTCTGAAATGATTCGTATACCAACCCAAAGCT 4116
QY 1401 ProGlyIleAspAspIleLysThrLeuGluGluLysLeuArgSerLeuPheSerGluHis 1420
Db 4117 CCTGAAATGATGACATTAAGACTCTGAAGAAAGTGGCGTCTGTTCAAGGAACAC 4176
QY 1421 SerSerSerGlyAlaGluHisAlaSerValSerLeuGluThrSerLeuValIleGluSer 1440
Db 4177 AGCTCATCTGAGCTCAGCATGCTCTGTCTCATCTGAGACCTCACTAGTACATAGAGAC 4236
QY 1441 ThrValThrProGlyIleProThrThrAlaValAlaProSerLysLeuLeuThrSerThr 1460
Db 4237 ACTGTCAACAGAGCATCCCACTACTGCTGTGTCACCAACAAACTCTGACTTCTAC 4296
QY 1461 ThrSerThrCysLeuProProThrAsnLeuProLeuGlyThrValAlaLeuProValThr 1480
Db 4297 ACAAGTACTGTCTTACCAACCAAAATTAACCATAGAACAGTGTGTCAGGTACCA 4356
QY 1481 ProValValThrProGlyGluValSerThrProValSerThrThrThrSerGlyValLys 1500
Db 4357 CCAAGTGTACACCTGGGCAAGTTTCTACCCAGTCAAGCACTACTCAATCAGAGTGA 4416
QY 1501 ProGlyThrAlaProSerLysProProLeuThrLysAlaProValLeuProValGlyThr 1520
Db 4417 CTTGAACTGCTCTCCCAAGCCACTTACTAAGGCTCCGATGCTGCCAGTGGTACT 4476
QY 1521 GluLeuProAlaGlyThrLeuProSerGluGluLeuProProPheProGlyProSerLeu 1540
Db 4477 GAACCTTCAGAGGTACTCTACCAAGAGAGGTGTCACCTTTTCCAGGACCTTCTCTA 4536
QY 1541 ThrGluSerGluGluProLeuGluAspLeuAspAlaGluLeuArgThrLeuSerPro 1560
Db 4537 ACCAGTCCCAAGCACTCTAGAGACTTATGTCAATGAGAAAGCACTTAGTCCA 4596
QY 1561 Glu**IleThrValThrSerAlaValGlyProValSerMetAlaAlaProThrAlaIle 1580
Db 4597 GAGATKATACAGAGACTTCTGCGGTGTGCTGTGTCTCAAGCGGCTCCAAACAGCAATC 4656
QY 1581 ThrGluAlaGlyThrGluProGluGlyValSerGluValLysGluGlyProValLeu 1600
Db 4657 ACAGAGAGAGAACAGCTTCAAGAGGTTTCTCAAGTCAAGAAAGGCGCTGTCTTA 4716
QY 1601 AlaThrSerSerGlyAlaGlyValPheLysMetGlyArgPheGluValSerValAlaAla 1620
Db 4717 GCAACTAGTTCAGAGAGCTGGTGTTTTAAGATGAGAGATTTACAGGTTTCTGTTGACGA 4776
QY 1621 AspGlyAlaGluGlyGlyLysAsnLysSerGluAspAlaLysSerValHisPheGlu 1640

Db 4777 GACGGTCCCAAGAAAGAGGTAAATAAGTCAGAAAGATGCAAAATCTGTTTCATTTTGA 4836
Qy 1641 SerSerThrSerGlnSerSerValLeuSerSerSerProGlnSerThrLeuVallys 1660
Db 4837 TCCACACCTCAGAGTCTCTCAGTCTATCAAGTATGATCCAGAGATCACTTGGTGAAA 4896
Qy 1661 ProGlnProAsnGlyIleThrIleProGlyIleSerSerAspValProGlnSerLahis 1680
Db 4897 CCAAGGCCGAATGGATAGCAATCCCTGATCTTTCAAGATGTCCAGAGATGCCAAC 4956
Qy 1661 LysThrThrAlaSerGlnAlaLysSerAspThrGlyGlnProThrLysValGlyArpHe 1700
Db 4957 AAAATACCTGCTCAGAGCAAAAGTCAGACACTGGGACCTACCAAGGTTGACATTTT 5016
Qy 1701 GlnValThrThrThrAlaAsnLysValGlyArgPheSerValSerLysThrGlnAspLys 1720
Db 5017 CAGGTGACAACTACGCAAAAGAGTGGTGGTCTCTGTATCAAAAACAGAGGACAG 5076
Qy 1721 IleThrAspThrLysLysGlnGlyProValAlaSerProProPheMetAspLeuGln 1740
Db 5077 ATCACTGACACAAAGAAAGAGGACAGTGGCATCTCTCTTATGATTTGAAACA 5136
Qy 1741 AlaValLeuProAlaValIleProLysLysGlnLysProGlnLeuSerGlnProSerHis 1760
Db 5137 GCTGTTCTTCTGCTGTATCAAAAGAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 5196
Qy 1761 LeuAsnGlyProSerSerSerProGlnAlaAlaPheLeuSerArgAspValAspAspGly 1780
Db 5197 CTAAATGGGCGCTCTTGACCCGAGGCGCTTTTAAAGTAAAGGAAAGTGGATGATAGT 5256
Qy 1781 SerGlySerProHisSerProHisGlnLeuSerSerLysSerLysProSerGlnAsnLeu 1800
Db 5257 TCCGGTATCCACACTGCCCCATCAGCTGAGCTAAAGAGCTTCCAGCCAGAAATCTA 5316
Qy 1801 SerGlnSerLeuSerAsnSerPheAsnSerSerLysMetSerSerAspAsnGlnSerAsp 1820
Db 5317 AGTCAAAGCCTTAGTAATTCATTTTAACTCTTTCATGAGTACGCAATGAGTACAGAT 5376
Qy 1821 IleGlnAspGlnAspLeuLysLeuGlnLeuArgLysLeuArgLysLeuLysLeuLysLeu 1840
Db 5377 ATCAAGATGAGAACTTAAAGTATGAGTGGACACATCAAGATAAACATCTCAAAAG 5436
Qy 1841 IleGlnAspLeuGlnSerArgGlnLysHisGlnIleGlnSerLysLysThrLysLeuGly 1860
Db 5437 ATTCAAGAGCTCGAGAGTCCGCAAGAGCATGAATGATCTTTGTATACCAACTGGGC 5496
Qy 1861 LysValProProAlaValIleIleProProAlaAlaProLeuSerGlyArgArgArgArg 1880
Db 5497 AAGGTGCCCCCTGCTGTTATTTATCCCACTGCTCCCTTTCAGGAGAAAGCAAGCA 5556
Qy 1881 ProThrLysSerLysGlnSerLysSerSerArgSerSerSerLysGlnLysLysSerPro 1900
Db 5557 CCACTTAAAGCAAAAGAGGAGCAAAATCTAGTCAAGCACTTCTTGGGAAATAAAGCCCC 5616
Qy 1901 GlnLeuSerGlyAsnLeuSerGlyGlnSerAlaAlaSerValLeuHisProGlnGlnThr 1920
Db 5617 CAGCTTTCAGGTAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5676
Qy 1921 LeuHisProProGlyAsnIleProGlnSerGlyGlnAsnGlnLeuGlnProLeuLys 1940
Db 5677 CTCACACCTCTCTGCAACATCCAGAGTCCGAGCAATCAGCTGTTACACCCCTTAAAG 5736
Qy 1941 ProSerProSerSerAspAsnLeuLysSerAlaPheThrSerAspGlyAlaIleSerVal 1960
Db 5737 CCACTCTCTCTCAAGTACCACTCTTATTCAGCTTTCACAGTGAAGGCGCATTTTCAGTA 5796
Qy 1961 ProSerLeuSerAlaProGlyGlnGlyThrSerSerThrAsnThrValGlyAlaThrVal 1980
Db 5797 CCAAGCCTTTCTGCTCAAGGTCAAGGAAACGAGCAACCACTGTTGGGGAACAGTGS 5856
Qy 1981 AsnSerGlnAlaAlaGlnAlaGlnProProAlaMetThrSerSerArgLysGlyThrPhe 2000

Db 5857 AACAGCCAGGCGCCCAAGCTCAGCTCTGCACTGATGAGTCCAGACAGAAAGGCACATTC 5916
Qy 2001 ThrAspAspLeuHisLysLeuValAspAsnThrAlaArgAspAlaMetAsnLeuSerGly 2020
Db 5917 ACAGATGACTTGCAACAAAGTGGTGTAGACAAATGGGCCCAAGATGCGCATGATCTCAAGC 5976
Qy 2021 ArgArgGlySerLysGlyHisMetAsnThrGlyGlyProGlyLysMetAlaArgLysPheSer 2040
Db 5977 AGGAGAGGAGGCAAGGAGCACTGATATTAAGGGCCCTGGAATGGCAAGAAAGTTCTCT 6036
Qy 2041 AlaProGlyGlnLeuCysIleSerMetThrSerAsnLeuGlyGlySerAlaProIleSer 2060
Db 6037 GCACCTGGGCAACTGCTGATCTCCATGATGACCTGAAACCTGGGCTCTGCCCCATCTCT 6096
Qy 2061 AlaAlaSerAlaThrSerLeuGlyHisPheThrLysSerMetCysProProGlnGlnThr 2080
Db 6097 GCAGATCAGTACTCTCTCTGATGACCTTCAACAAAGTATGATGCCCCCAAGAGAT 6156
Qy 2081 GlyPheProAlaThrProPheGlyAlaGlnThrPheSerGlyThrGlyProAlaProGln 2100
Db 6157 GGCCTTCCAGCTACCCCATTTGGCGCTCATGAGTGGAGCGGTGGCCCAAGACCAAG 6216
Qy 2101 ProLeuGlyGlnPheGlnProValGlyThrAlaSerLeuGlnAsnPheAsnLysSerAsn 2120
Db 6217 CCACCTGGCCAGTTTCAACCTGTGGAACTGCTCTTGCAGAAATTTCAACATCAGCAAT 6276
Qy 2121 LeuGlnLysSerLysSerAsnProProGlySerAsnLeuArgThrThr 2136
Db 6277 TTGCAAGAAATCCATCAGCAACCCCAAGGCTTCAACCTGCGAGCACT 6324
RESULT 7
ID ADC99101 standard; cDNA; 7152 BP.
XX
AC ADC99101;
XX
DT 01-JAN-2004 (first entry)
XX
DE Human KRP cDNA - SEQ ID 54.
XX
KW anti-HIV; antiallergic; antiinflammatory; antianaemic; antiparkinsonian;
KW nootropic; anticonvulsant; antiarteriosclerotic; antiasthmatic;
KW immunosuppressive; antithyroid; cytostatic; hepatocytologic; dermatologic;
KW antidiabetic; nephrotoxic; antitumor; thymic; neuroprotective;
KW osteopathic; antirheumatic; antiparasitic; antihelminthic; antispasmodic;
KW uterine; ophthalmologic; fungicide; antirheumatic; haemostatic; KRP;
KW cancer; developmental disorder; atherosclerosis; cirrhosis; hepatitis;
KW Alzheimer's disease; Parkinson's; autoimmune; inflammatory; Crohn's;
KW diabetes mellitus; viral; bacterial; fungal; parasitic; protozoan;
KW helminthic infection; transgenic; gene therapy; human; ss; gene.
XX
OS Homo sapiens.
XX
PN WO2003033680-A2.
XX
PD 24-APR-2003.
XX
PF 17-OCT-2002; 2002WO-US033723.
XX
PR 19-OCT-2001; 2001US-0345474P.
XX
PR 12-NOV-2001; 2001US-0343910P.
XX
PR 13-NOV-2001; 2001US-0333098P.
XX
PR 16-NOV-2001; 2001US-0332424P.
XX
PR 30-NOV-2001; 2001US-0334288P.
XX
PA (INCY-) INCYTE GENOMICS INC.
XX
PI Bandman O, Baughn MR, Becha SD, Borowsky ML, Duggan BM,
PI Emerling BM, Foreyche IU, Gandhi AR, Gorvad AE, Griffin JA,
PI Gururajan R, Hafalia AJA, Khan FA, Lal PG, Lee EA, Lee SY,
PI Lindquist EA, Lu DAM, Lu Y, Marquis JP, Nguyen DB, Arvizu CS;

PI Rankumar J, Recipon SA, Richardson TW, Swarnakar A, Tang YT,
PI Thornton MB, Tran UK, Chawla NK, Warren BA, Yang J, Yao MG, Yue H;
PI Zebajadian Y;
XX
XX WPI: 2003-403214/38.
DR P-PSDB; ADC99049.

XX New human kinases and phosphatases and polynucleotides, useful for
PT diagnosing, treating or preventing autoimmune or inflammatory disorders
PT (e.g. AIDS, allergy or anemia), multiple sclerosis, osteoarthritis,
PT cancer or hepatitis.

XX Claim 5; SEQ ID NO 54; 424pp; English.

XX The invention relates to a novel isolated polypeptide which is a human
CC kinase and phosphatase (KPP). The KPP polypeptides, polynucleotides,
CC agonists and antagonists are useful for diagnosing, treating or
CC preventing cell proliferative disorders such as atherosclerosis,
CC cirrhosis, hepatitis and cancer, developmental disorders e.g. mental
CC retardation, neurological disorders including Alzheimer's disease and
CC Parkinson's disease, autoimmune and inflammatory disorders such as
CC Crohn's disease and diabetes mellitus and finally, viral, bacterial,
CC fungal, parasitic, protozoan or helminthic infections. Furthermore, the
CC polynucleotides encoding KPP may be useful for creating transgenic
CC animals to model human disease, as well as during gene therapy
CC procedures. The current sequence is that of the human KPP cDNA of the
CC invention.

XX Sequence 7152 BP; 1915 A; 1952 C; 1682 G; 1603 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	Length:	7152
Score: 0	Matches:	2107
Percent Similarity: 10646.00	Conservative:	0
Best Local Similarity: 98.64%	Mismatches:	1
Query Match: 98.46%	Indels:	28
DB:	Gaps:	1

US-10-010-720-14 (1-2136) X ADC99101 (1-7152)

QY	1	MetSerGlyGlyAlaAlaGlnLysGlnSerThrProGlySerLeuPheLeuSerPro	20
DB	361	ATGCTGGCGGCGCGCGAGAACGAGACGACTCCCGGTTCCCTGTTCTCCGCGG	420
QY	21	ProAlaProAlaProLysAsnGlySerSerSerSerSerValGlyGlnLysLeuGly	40
DB	421	CCGGCTCTGCGCCCGAGAAATGGCTCCAGCTCCGATCTCTCGTGGGGAAGAACTGGA	480
QY	41	AlaAlaAlaAlaAspAlaValThrGlyArgThrGlnGlnLysArgArgArgHisThr	60
DB	481	GCCGGCGCGCGCGAGCTGTGACCGCGGAGACCGAGAGTACAGCGCGCGCACACT	540
QY	61	MetAspLysAspSerArgGlyAlaAlaAlaThrThrThrThrGlnHisArgPhePhe	80
DB	541	ATGACAAAGACAGCCGCGGCGCGCGCGAGCACTACACACAGACACACGCTTCTTC	600
QY	81	ArgArgSerValIleCysAspSerAsnAlaThrAlaLeuGlnLysProGlyLeuProLeu	100
DB	601	CGCGGAGCGCTCATCTGCACTCCATGCTCCAGCTGAGACTTCCGCGCTTCTCTT	660
QY	101	SerLeuProGlnProSerIleProAlaAlaValProGlnSerAlaProGlnProHis	120
DB	661	TCCCTGCGCCGAGCCAGCATCCCGCGGCTGTCCCGGAGAGTGTCCACCGGAGCCAC	720
QY	121	ArgGlnGlnThrValThrAlaThrAlaThrSerGlnValAlaGlnGlnProProAlaAla	140
DB	721	CGGAAAGAGACCGGAGCCGACCGCTTCCAGGTAGCCACAGACCTCCAGCCCT	780
QY	141	AlaAlaProGlyGlnGlnAlaValAlaGlyProAlaProSerThrValProSerSerThr	160
DB	781	GCCGCGCTCGGGAACAGCGCGCTCGCGGCGCTTCCGCTGACTGTCCACGACGATACC	840
QY	161	SerLysAspArgProValSerGlnProSerLeuValGlySerLysGlnGlnProPro	180

DB	841	AGCAAAACCGCCAGAGTCCAGCTTGTGGGAGCAAAAGAGCGCCCGCGG	900
QY	181	AlaArgSerGlySerGlyGlySerAlaAlaGlnProGlnGlnLysSerGlnGln	200
DB	901	GCGAAGAGTGCAGCGCGCGCGGAGCCAGACGACGACGAGAAAGGACGACGAG	960
QY	201	GlnAspAspIleGlnGlnLysGlnThrAlaValAlaGlyMetSerAsnAspGlyArgPhe	220
DB	961	CAGATATATTCGAAGAGCTGAGAGCCAGGCGGTGGGANTGTAAAGATGCGCGCTT	1020
QY	221	LeuLysPheAspIleGlnIleGlyArgGlySerPheLysThrValTyrLysGlyLeuAsp	240
DB	1021	CTCAAGTTTGACATCGAAATCGCAGAGCTCTTTAAGACGCTCAAAAGGCTCGAC	1080
QY	241	ThrGlnThrThrValGlnValAlaThrCysGlnLysGlnAspArgLysLeuThrLysSer	260
DB	1081	ACTGAAACCAACCGGAGAGTCCGTGTGAACTGCAAGATCGAAATTTAAACAAAGCT	1140
QY	261	GlnArgGlnArgPheLysGlnGlnAlaGlnMetLeuLysGlyLeuGlnHisProAsnIle	280
DB	1141	GAGAGGAGAGATTTTAAAGAAAGCTGAATGTTTAAAGGCTTCAGCATCCCAATAT	1200
QY	281	ValArgPheTyrAspSerTyrGlySerThrValLysGlyLysLysCysIleValLeuVal	300
DB	1201	GTTAGATTATATGATTCCTGGGAATCCACAGTAAAGAAAGAGTGCATGTTTGTG	1260
QY	301	ThrGlnLysMetThrSerGlyThrLeuLysThrTyrLeuLysArgPheLysValMetLys	320
DB	1261	ACTGAACCTTATGAGCTGTGAACACTTAAACGATCTGAAAGGTTTAAAGTATGAG	1320
QY	321	IleLysValLeuArgSerTyrCysArgGlnIleLeuLysGlyLeuGlnPheLeuHisThr	340
DB	1321	ATCAAAATTTAAGAACTGGTGGCGGCAATCTTAAAGGCTCTTCAAGTTCTTCAATCT	1380
QY	341	ArgThrProProIleIleHisArgAspLeuLysCysAspAsnIlePheIleThrGlyPro	360
DB	1381	CGAATCCACCTATCATTCACCGGATCTTAAATGTACAAACATTTATACCGGCGCT	1440
QY	361	ThrGlySerValLysIleGlyAspLeuGlnLeuAlaThrLeuLysArgAlaSerPheAla	380
DB	1441	ACTGGCTGACGAAAGATTGAGACTCGGTGTGCAACCTGAAAGCGGCTTCTTTTCC	1500
QY	381	LysSerValIleGlyThrProGlnPheMetAlaProGlnMetTyrGlnGlnLysTyrAsp	400
DB	1501	AAGAGTGTGATGATGATCCAGAGTTCTATGGCCCTGAGATGATGAGAAATATGAT	1560
QY	401	GlnSerValAspValTyrAlaPheGlyMetCysMetLeuGlnMetAlaThrSerGlyTyr	420
DB	1561	GAATCCGTTGACGTTTATGCTTTTGGATGTGATGCTTGAGATGGCTACATCGAAAT	1620
QY	421	ProTyrSerGlnCysGlnAsnAlaAlaGlnIleTyrArgArgValThrSerGlyValLys	440
DB	1621	CTTAACTGAGATGCCAAATGCTGCGAGATTCACGTCGCGAGACGAGGGGTGAAG	1680
QY	441	ProLysSerPheAspLysValAlaIleProGlnValLysGlnIleGlnGlyCysIle	460
DB	1681	CCAGCCAGTTTGAACAAGTACATCTCGAATGAGAAATTAATTGAAGATGACATA	1740
QY	461	ArgGlnAsnLysAspGlnArgTyrSerIleLysAspLeuLeuAsnHisAlaPhePheGln	480
DB	1741	CGACAAAACAAAGATGAAGATATTCATCAAAACCTTTGAAACATGCTTCTTCCAA	1800
QY	481	GlnGlnThrGlyValArgValGlnLeuAlaGlnGlnAspArgLysGlnLysIleAlaIle	500
DB	1801	GAGGAACAGAGATGAGGTGAAATGACAGAAAGATGATGAGAAATTAATGCCATA	1860
QY	501	LysLeuThrPheLysArgIleGlnAspIleLysLysLeuLysGlyTyrTyrAspAsnGln	520
DB	1861	AAATTTAGGCTAGCATATTGAAGATATTAAAGAAATTAAGGGAATAATCAATGAA	1920
QY	521	AlaIleGlnPheSerPheAspLeuGlnArgAspValProGlnAspValAlaGlnGlnMet	540

Db 1921 GCTATTGAGTTTCTTTGATTAGAGAGATGTCCTCAAGAGATGTTGCACAGAAATG 1980
 QY ValGluSerGlyThrValCysGluGluValAspHisIleThrMetAlaIleValAsp 560
 Db 1991 GTAGAGTCTGGGTATGCTGTGTAAGATGATCAAGACCATGGCTTAAGCTATCAAAAC 2040
 QY ArgValSerLeuIleLeuArgValAspGluValArgGluValArgGluGluGlu 580
 Db 2041 AGAGTATCATTTATTAAGAGAAACGAGCGCGGAGTTGTTGTTGCGGAGAGAGAA 2100
 QY LysLysLysGluGluGluSerLeuLysGluGluValGluGluSerLeuAspGln 600
 Db 2101 AAAAAAGCAGGAGAGAGAGAGCTCAACAGCGAGTGAACATCCAGTCTCCAG 2160
 QY ThrGlyIleLysGluLeuProSerAlaSerThrGlyIleProThrAlaSerThrThr 620
 Db 2161 ACAGGAATCAAGACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCA 2220
 QY AlaSerValSerThrGluValGluProGluGluProGluAlaAspGlnHisGluGlnLeu 640
 Db 2221 GCTTCAGTTTCTACACAGTGAAGACCTGAAGACCTGAAGGAGATCAACATCAACACTA 2280
 QY GlnThrGluGlnProSerIleSerValLeuSerAspGlyThrValAspSerGlyGlnGly 660
 Db 2281 CAGTACCGACCAACCGATATCTGTGTATCTGATGGACGCTTGACAGTGTCAAGGA 2340
 QY SerSerValPheThrGluSerArgValSerSerGluGlnThrValSerThrGlySerGln 680
 Db 2341 TCCCTGCTTCAACAGATCTGAGTGAAGACCAACAGACAGTTCATTAAGTTCCTCA 2400
 QY HisGluGlnAlaHisSerThrGlyThrValProGlyHisIleProSerThrValGlnAla 700
 Db 2401 CATGAACAGGACATTTCTACAGGACAGTCCAGGCGATATACCTTACTGCTCCAAACA 2460
 QY GlnSerGlnProHisGlyValTyrProProSerSerValAlaGlnGlyGlnSerGlnGly 720
 Db 2461 CAGTCTCAGCCCGCATGGGGTATCCACCTCAAGTGTG----- 2499
 QY GlnProSerSerSerSerLeuThrGlyValSerSerSerGlnProIleGlnHisProGln 740
 Db 2499 ----- 2499
 QY GlnGlnGlnGlnGlyIleGlnGlnThrAlaProProGlnGlnThrValGlnTyrSerLeuSer 760
 Db 2500 ---CAGCAGGGAATACAGCAGACAGCCCTCCTCAACAGACAGTGAATATCACTTCA 2556
 QY GlnThrSerThrThrSerSerGluAlaThrThrAlaGlnProValSerGlnProGlnAlaPro 780
 Db 2557 CAGACATCAACCTCAAGTGAAGGCCACTGACACAGCCAGTGAAGTCAACCTCAAGCTCCA 2616
 QY GlnValLeuProGlnValSerAlaGlyLysGlnSerThrGlnGlyValSerGlnValAla 800
 Db 2617 CAATCTTGGCTCAAGTATCAAGCTGGAAGAACAGTACTCAAGGAGTCTCTCAAGTGTCT 2676
 QY ProAlaGluProValAlaValAlaGlnProGlnAlaThrGlnProThrThrLeuAlaSer 820
 Db 2677 CCTGCAGAGCGCAGTTCAGTGAAGCAGCCCAAGCTACCCAGCCAGCACCTTGGCTTCC 2736
 QY SerValAspSerAlaHisSerAspValAlaSerGlyMetSerAspGlyAsnGluAsnVal 840
 Db 2737 TCTGTAAACAGTGAACATTCAGATGTGTCTCAGGTATGAGTGAATGCAATGAAGAACGTC 2796
 QY ProSerSerSerGlyArgHisGluGlyValArgThrThrLysArgHisGlyArgLysSerVal 860
 Db 2797 CCATCTTCCAGTGAAGGACATGAAGAGAACTCAAAACGCGATTACCGAAAAATCTGTA 2856
 QY ArgSerArgSerArgHisGluLysThrSerArgProLysLeuArgIleLeuAsnValSer 880
 Db 2857 AGAGTGTGCTCTCGACATGAAAAAACTTCAAGCCCAAAATTAAGATTTTGAATGTTCA 2916
 QY AsnLysGlyAspArgValAlaGluCysGlnLeuGluThrHisAsnArgLysMetValThr 900
 Db 2917 AATTAAGAGACCGAGTATGATGATGCTCAATTAGAGACTCATTAATAGAAAAATGTTTCA 2976

QY PheLysPheAspLeuAspGlyAspAsnProGluGluIleAlaThrIleMetValAsnAsn 920
 Db 2977 TTCAAAATTTGACCTAGATGATGACAAACCCAGAGAGATGCAACATTAATGATGAACAAT 3036
 QY AspPheIleLeuAlaIleGluArgGluSerPheValAspGlnValArgGluIleIleGlu 940
 Db 3037 GACTTATTTACGATATGAGAGAGAGATGCTTTGTGGATTCAGATGCGAGAAATTAATGAA 3096
 QY LysAlaAspGluMetLeuSerGluAspValSerValGluProGluGlyAspGlnGlyLeu 960
 Db 3097 AAAGCTATGAAATGCTCAGTGAAGATGCTAGTGAACCAAGAGGATGATCAGGAGATTG 3156
 QY GluSerLeuGlnGlyLysAspAspTyrGlyPheSerGlySerGlnLysLeuGluGlyGlu 980
 Db 3157 GAGAGTCTACAGAGAAAGAGATGACTATGCTTTTCAAGGTTCTCAAGAAATGGAAGAGAG 3216
 QY PheLysGlnProIleProAlaSerSerMetProGlnGlnIleGlyIleProThrSerSer 1000
 Db 3217 TTCAACGACCAATTCCTGCGCTTCCATGCGCAACGAAATAGGCATTCCTACCACTTCT 3276
 QY LeuThrGlnValValHisSerAlaGlyArgArgPheIleValSerProValProGluSer 1020
 Db 3277 TTAACCTCAAGTGTTCATTTCTGCGGAAAGCGGTTTATAGAGATCCTGTCAGAAAGC 3336
 QY ArgLeuArgGluSerLysValPheProSerGluIleThrAspThrValAlaAlaSerThr 1040
 Db 3337 CGATTTCAGAGATCAAAAGATTTTCCAGTGAATTAACAGTATACAGTGTGCTGCTTCA 3396
 QY AlaGlnSerProGlyMetAsnLeuSerHisSerAlaSerSerLeuSerLeuGlnGlnAla 1060
 Db 3397 GCTCAGAGCCCTGGAATGAATTCCTGCTCAGCTGCAATCCTTACTGCTTCAACAGGCC 3456
 QY PheSerGluLeuArgArgAlaGlnMetThrGlnGlyProAsnThrAlaProProAsnPhe 1080
 Db 3457 TTTTCTCAACTTAACAGCTGCCCCAATGACAGAAAGACCAATACAGCACTCCAAACTTT 3516
 QY SerHisThrGlyProThrPheProValValProPheLeuSerSerIleAlaGlyVal 1100
 Db 3517 AGTCATACAGGACCAACATTTCCAGTGAATGCTCTCTTTCTTAATATACATTTGCTGAGTC 3576
 QY ProThrThrAlaAlaAlaThrAlaProValProAlaThrSerSerProProAsnAspIle 1120
 Db 3577 CCAACCAAGAGAGAGCAACAGCCAGCTCCTGCAACAGCAGCCCTCTCAATGACATT 3636
 QY SerThrSerValIleGlnSerGluValThrValProThrGluGluGlyIleAlaGlyVal 1140
 Db 3637 TCCACATCAGTAATTCAGTCTGAGGTACAGTGCCTCACTGAAGAGGGGATGCTGAGATT 3696
 QY AlaThrSerThrGlyValValThrSerGlyLysLeuProIleProProValSerGluSer 1160
 Db 3697 GCCACGACCAAGGTGTGGAATCTTCAAGTGTGTCCCATACCACTGATGTGGAATCA 3756
 QY ProValLeuSerSerValValSerSerIleThrIleProAlaValValSerIleSerThr 1180
 Db 3757 CCAAGTATCTTCCAGCGATGTTCAAGTATCAACATACCTGCAAGTGTCTCAATCTACT 3816
 QY ThrSerProSerLeuGlnValProThrSerThrSerGluIleValValSerSerThrAla 1200
 Db 3817 ACATCCCGCTCAGTCAAGTCCCAACATCAACATCAAGATCGTTGTTTCAATGACAGA 3876
 QY LeuThrProSerValThrValSerAlaThrSerAlaSerAlaGlyLysThrThrAlaThr 1220
 Db 3877 CTGATCTCTTCAATGATGATTTCAAGCACTTCACTCTGAGGGGAGCACTGCTGATCC 3936
 QY ProGlyProLysProProAlaValValSerGlnGlnAlaAlaGlySerThrThrValGly 1240
 Db 3937 CCAAGTCTTAAGCTCAGCGTATGATTCAGAGGACCAAGCACTACCTGATGGA 3996
 QY AlaThrLeuThrSerValSerThrThrThrSerPheProSerThrAlaSerGlnLeuSer 1260
 Db 3997 GCCACATTAACATCAGTTCTACACCACTTCACTTCCCAAGCAAGCTTCAAGCTGTCC 4056

QY 1261 ILeuInleuSerSerSerThrThrProThrLeuAlaGluThrValValSerAla 1280
DB 4057 ATTCAAGCTTAGACACAGAGTACTTCTACTCTTACTAGTGAACCGTGTAGTTAGCGCA 4116
QY 1281 HisSerLeuAspLysThrSerHisSerSerThrThrGlyLeuAlaPheSerLeuSerAla 1300
DB 4117 CACCTACATAGTAAGACATCTCATAGCAAGTAACTGGATTGGCTTCTCCCTCTCGCA 4176
QY 1301 ProSerSerSerSerSerProGlyValAlaGlyValSerSerThrIleSerGlnProGlyGly 1320
DB 4177 CCACTCTCTCTCTCTCTCTCGAGACAGAGTGTCTGATTATTTCTCATGCTCGTGTGG 4236
QY 1321 LeuHisProLeuValIleProSerValIleAlaSerThrProIleLeuProGlnAlaAla 1340
DB 4237 CTGATCTCTTGGTTCATTTCCATTCAGTAGTACTTCTACTCTCTATTTCTTCCCAAGCA 4296
QY 1341 GlyProThrSerThrProLeuLeuProGlnValProSerIleProProLeuValGlnPro 1360
DB 4297 GGACCTCTCTTACACCTTTATTTACCCCACTACTGTTATCCCACTTGTGTACAGCT 4356
QY 1361 ValAlaSerValProAlaValAlaGlnThrLeuIleHisSerGlnProGlnProAlaLeu 1380
DB 4357 GTTCCAAATGTGCTGCTGTACAGACAGACTAATTCATGTCAGCTCAACCAAGCTTTG 4416
QY 1381 LeuProAsnGlnProHisSerHisSerProGlyValAlaSerAspThrGlnProIleVala 1400
DB 4417 CTTCACCAACCAAGCCCACTACTGTTGCTTGAAGATTCTGATACCAACCAAGCT 4476
QY 1401 ProGlyIleAspAspIleLysThrLeuGlnGlyLysLeuArgSerLeuPheSerGluHis 1420
DB 4477 CCTGGAATTGATGACATTAAGACCTCTAGAAAGAAAGCTGCGCTCTGTTCAAGTAA 4536
QY 1421 SerSerSerGlyAlaGlnHisAlaSerValSerLeuGlnThrSerLeuValIleGluSer 1440
DB 4537 AGCTCATCTGAGCTCAGACATGCTCTGTCTACTGAGACCTCACTAGTATAGAGGC 4596
QY 1441 ThrValThrProGlyIleProThrThrAlaValAlaProSerLysLeuLeuThrSerThr 1460
DB 4597 ACTGTACACCAAGGATCCCACTACTGCTGTGTGACCAACCAACTCTTCACTTACC 4656
QY 1461 ThrSerThrCysLeuProProThrAsnLeuProLeuGlyThrValAlaLeuProValThr 1480
DB 4657 ACAAGTACTTCTTACCAACCAATTTACCACTAGGACAGATGCTTTCACAGTACA 4716
QY 1481 ProValValThrProGlyGlnValSerThrProValSerThrThrThrSerGlyValLys 1500
DB 4717 CCAGTGGTCACACCTGGGCAAGTTCTACCCAGTCACACTACTACATCCAGAGTGA 4776
QY 1501 ProGlyThrAlaProSerLysProProLeuThrLysAlaProValLeuProValGlyThr 1520
DB 4777 CTTGGAACTGCTCCTCCCAAGCACTCTTAAGCTCCGGTGTCCAGTGGTACT 4836
QY 1521 GluLeuProAlaGlyThrLeuProSerGlnGlnLeuProProPheProGlyProSerLeu 1540
DB 4837 GAACCTTCAGACGAGTACTCTACCCAGCAGACGCTCCATCTTTCAGAGACTTCTCTA 4896
QY 1541 ThrGlnSerGlnGlnProLeuGlnAspLeuAspAlaGlnLeuArgArgThrLeuSerPro 1560
DB 4897 ACCAGTCCACAGCACTCTAGAGATCTTGAATGCTCAATTAGAGAAACCTTAGTCA 4956
QY 1561 Glu**IleThrValThrSerAlaValGlyProValSerMetAlaAlaProThrAlaIle 1580
DB 4957 GAGATTATCACAGTACTTTCGGGTGTGTCTGTCTCATGCGGGCTCCAAAGCACTC 5016
QY 1581 ThrGlnAlaGlyThrGlnProGlnLysGlyValSerGlnValLysGlnGlyProValLeu 1600
DB 5017 ACAGAAAGCAAGAAACACACTCAGAAAGGTGTTCTCAAGTCAAGAAAGAGCCCTGCTCA 5076
QY 1601 AlaThrSerSerGlyAlaGlyValAlaPheLysMetGlyArgPheGlnValSerValAlaAla 1620
DB 5077 GCAACTAGTTCAGAGCGTGTGTGTTTAAAGTGGACATTTCAAGATTTCTGTTCACGA 5136
QY 1621 AspGlyAlaGlnLysGlnGlyLysAsnLysSerGluAspAlaLysSerValHisPheGlu 1640
DB 5137 GACGTGCCCCAGAAAGGGGTAAATAATCTCAGAAATCCAAAGTGTCTCATTTGAA 5196
QY 1641 SerSerThrSerGlnSerSerValLeuSerSerSerSerProGlnSerThrLeuValLys 1660
DB 5197 TCCAGACCTCCAGAGTCTTAGTCTATCAAGTATGATCCAGAGATCACTTGTGTGA 5256
QY 1661 ProGluProAsnGlyIleThrIleProGlyIleSerSerAspValProGlnSerAlaHis 1680
DB 5257 CCAGAGCCGATAGGCATACCATCTCTGATATCTTCAATGTGCCAGAGAGTCCAC 5316
QY 1681 LysThrThrAlaSerGlnAlaLysSerAspThrGlyGlnProThrLysValGlyArgPhe 1700
DB 5317 AAAACTACTGCTCAGAGCAAAAGTACAGACTGGGACGCTTCAAGAGTTGACGTTT 5376
QY 1701 GlnValThrThrThrAlaAsnLysValGlyArgPheSerValSerLysThrGlnAspLys 1720
DB 5377 CAGGTGCAACTAGCAAAACAAAGTGGTGTGTTCTGTATCAAAACCTGAGGACAA 5436
QY 1721 IleThrAspThrLysLysGlnGlyProValAlaSerProPheMetAspLeuGln 1740
DB 5437 ATCACTGACAAAGAAAGAAAGAGCACGTGGCATCTCTCTTTATGATTTGGAA 5496
QY 1741 AlaValLeuProAlaValIleProLysLysGlnLysProGluLeuSerGluProSerHis 1760
DB 5497 GCTGTCTTCTCTGCTGTGATACCAAGAAAGAGAGCTGAACGTCAAGAGCTTCACAT 5556
QY 1761 LeuAsnGlyProSerSerAspProGlnAlaAlaPheLeuSerArgAspValAspAspGly 1780
DB 5557 CTAAATGGGCGCTTCTGACCCGAGAGCGCTTTTAAAGTAGAGATGATGATG 5616
QY 1781 SerLysProHisSerProHisGlnLeuSerSerLysSerLeuProSerGlnAsnLeu 1800
DB 5617 TCCGTATGTCACACTGCCCCATCACTAGCTCAAGAGCCTTCTTACGACAGATCTA 5676
QY 1801 SerGlnSerLeuSerAsnSerPheAsnSerSerLysSerSerAspAsnGluSerAsp 1820
DB 5677 AGTCAAAAGCTTAGTATTTATTTAACTCTTACCTGATGATGATGATGATGAT 5736
QY 1821 IleGlnAspGlnAspLeuLysLeuGlnLeuArgProLeuArgPheLysLeuGln 1840
DB 5737 ATCGAAGATGAAGCTTAAGTATGAGCTGAGACACTACCAAGTAAACATCTCAAG 5796
QY 1841 IleGlnAspLeuGlnSerArgGlnLysHisGlnIleGlnSerLeuThrLysLeuGly 1860
DB 5797 ATTGAGACCTCGAGAGTGGCAGAGATGAATTTGATTTGTATACCAACTGGGC 5856
QY 1861 LysValProProAlaValIleIleProProAlaAlaProLeuSerGlyArgArgArg 1880
DB 5857 AAGGTGCCCCCTGCTGTTATTTATCCCCAGCTCTCTCTTTCAGGAGAAAGCA 5916
QY 1881 ProThrLysSerLysGlySerLysSerSerArgSerSerSerLeuGlyAsnLysSerPro 1900
DB 5917 CCACTAAAGCAAAAGGAGCAATCTAGTCAAGCAAGTCTCTTGGGAAATAAAGCCCC 5976
QY 1901 GlnLeuSerGlyAsnLeuSerGlyGlnSerAlaAlaSerValLeuHisProGlnGlnThr 1920
DB 5977 CAGCTTTCAGTAACTGTCTGTCTGATCAGAGTCACTTCACTTCAACCCCAAGACC 6036
QY 1921 LeuHisProProGlyAsnIleProGlnSerGlyGlnGlnGlnLeuLeuGlnProLeuLys 1940
DB 6037 CTCACACCTCTGGCAATCCAGAGTCCGGGAGATTCAGCTGTTTACACCCCTTAA 6096
QY 1941 ProSerProSerSerAspAsnLeuThrSerAlaPheThrSerAspGlyAlaAlaSerVal 1960
DB 6097 CCACTCTCTCCAGTACACTTATTTACACTTCAACAGTATGATGATGATTTCACTA 6156
QY 1961 ProSerLeuSerAlaProGlyGlnGlyThrSerSerThrAsnThrValGlyAlaThrVal 1980
DB 6157 CCAAGCCTTTGCTCCAGGTCAGGAAAGCAAGCAAGCAAACTGTTGGGGCAACAGTG 6216
QY 1981 AsnSerGlnAlaAlaGlnAlaGlnProProAlaMetThrSerSerArgLysGlyThrPhe 2000

Db 901 ACTGAACCTATGAGCTCTGAGACCTTAAACGATCTATCTGAAAGGTTTAAAGTATGAG 960
 Qy 321 HLeuValLeuArgSerTPCyArgGlnHLeuLeuGlyLeuGlnPheLeuHisThr 340
 Db 961 ATCAAGGTTCTMAAGACTGGTCCGTCAGATCTTAAAGGCTCTTCAGTTCTTCTCACT 1020
 Qy 341 ArgThrProProlLeuHisArgAspLeuLeuCyAspAsnHLeuHLeuHLeuGlyPro 360
 Db 1021 CGAACTCAGCTATCATTCACCCGATCTTAAATGTGACAAACATCTTATCACCGGCCCT 1080
 Qy 361 ThrGlySerValLeuHLeuGlyAspLeuGlyLeuHLeuLeuGlyAspGlnAspPheAla 380
 Db 1081 ACTGGCTCAGTCAAGATTGAGAGCTGCTGCTGCAACCTGAAAGCGGCTCTTTTGGCC 1140
 Qy 381 LysSerValHLeuGlyThrProGlnPheMetAlaProGlnMetTyrGlnGlnLysTyrAsp 400
 Db 1141 AAGGTGTGATGATGATCCCGAGATGATGATGATGATGATGATGATGATGATGATGAT 1200
 Qy 401 GluSerValAspValTyrAlaPheGlyMetCysMetLeuGlnMetAlaThrSerGlyTyr 420
 Db 1201 GAATCCGTTGACGTTTATGCTTTGGGATGTCATGCTGATGATGATGATGATGATGAT 1260
 Qy 421 ProTyrSerGlnCysGlnAsnAlaAlaGlnHLeuTyrArgArgValThrSerGlyValLys 440
 Db 1261 CCTTACTCGAGTCCCAAAATGCTGCAAGATCTTACCGCTCCGTCGACCAAGTGGGTGAG 1320
 Qy 441 ProHLeuPheAspLysValAlaHLeuProGlnValLysGlnHLeuGlyCysHLeu 460
 Db 1321 CCAGCCGTTTGGCAAAAGTGAATCTCTGAGTGAAGGAAATTTATGAAAGATGCTATA 1380
 Qy 461 ArgGlnAsnLysAspGlnArgTyrSerHLeuAspLeuLeuAsnHLeuAlaPhePheGln 480
 Db 1381 CGACAAACCAAGATGAAGATATTCATCAAGACCTTTGAACCATGCTTCTTCCAA 1440
 Qy 481 GluGlnThrGlyValArgValGlnLeuHLeuGlnLysAspArgGlyGlnLysHLeuAla 500
 Db 1441 GAGAAACAGAGATACGGGTGAATTCAGAAAGATGATGAGAAATTAATGAGCCATA 1500
 Qy 501 LysLeuThrPheLeuArgHLeuAspHLeuLysLeuLysGlyLysTyrLysAspAsnGln 520
 Db 1501 AAATTTATGCTACGATTAAGATATTAAGAAATTAAGGAAATTAACAAAGATATGAA 1560
 Qy 521 AlaHLeuGlnPheSerPheAspLeuGlnArgAspValProGlnAspValAlaGlnLysMet 540
 Db 1561 GCTATGAGTTTCTTTGATTTAGAGAGATGCTCCAGAAAGATGTTGCAAGAAATG 1620
 Qy 541 ValGlnSerGlyTyrValCysGlnGlyAspHLeuLysThrMetAlaLysValHLeuAsp 560
 Db 1621 GTAGAGTCTGGGTATGCTCTGAGAGGTGATCAAGAACCATGCGCTAAAGCTATCAAAAG 1680
 Qy 561 ArgValSerLeuHLeuLysArgLysArgGlnArgGlnArgGlnArgGlnArgGlnArg 580
 Db 1681 AGAGATCATTAATTAAGAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAA 1740
 Qy 581 LysLysLysGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 600
 Db 1741 AAAAAAAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1800
 Qy 601 ThrGlyHLeuGlnLeuProSerAlaSerThrGlyHLeuProThrAlaSerThrHisSer 620
 Db 1801 ACAGAAATCAAGACACTCTCTCTCTGACACCGGACATACCTACTGCTTCAACCTTCA 1860
 Qy 621 AlaSerValSerThrGlnValGlnProGlnGlnProGlnAlaAspGlnHLeuGlnLeu 640
 Db 1861 GCTTAGTTTCTACCAAGATGAACTGAGAACTGAGGAGAGAGAGAGAGAGAGAGAGAG 1920
 Qy 641 GlnTyrGlnGlnProSerHLeuSerValLeuSerAspGlyThrValAspSerGlyGlnGly 660
 Db 1921 CAGTACCAAGCAACCAAGATATCTGTGTTATCTGATGGAGCGTTGAGAGTGTGAGGGA 1980
 Qy 661 SerSerValPheThrGlnSerArgValSerSerGlnGlnThrValSerTyrGlySerGln 680
 Db 1981 TCCTCTGCTTCAAGAACTCTGAGATGAGCAGCAACAGAGTTTCATATGATGCCCA 2040

Qy 681 HisGlnGlnAlaHisSerThrGlyThrValProGlnHisHLeuProSerThrValGlnAla 700
 Db 2041 CATGAACAGGACATTTCTACAGGACAGTCCAGGAGATATCTTACTTACTGTCAGCA 2100
 Qy 701 GlnSerGlnProHLeuGlyValTyrProProSerSerValAlaGlnGlnGlnSerGlnGly 720
 Db 2101 CAGTCTCAGCCCATGGGTATATCCACCTCAAGTGTG----- 2139
 Qy 721 GlnProSerSerSerSerLeuThrGlyValSerSerSerGlnProHLeuGlnHisProGln 740
 Db 2139 ----- 2139
 Qy 741 GlnGlnGlnGlyHLeuGlnGlnThrAlaProProGlnGlnThrValGlnTyrSerLeuSer 760
 Db 2140 ---CAGCAGGGAATACAGCAGACAGCCCTCTCTCAACAGACAGTGAATTCACCTTCA 2196
 Qy 761 GlnThrSerThrSerSerGlnAlaThrThrAlaGlnProValSerGlnProGlnAlaPro 780
 Db 2197 CAGACATCAACCTCCAGTGAAGCCACATGACAGCCAGTGAAGTCAAGCTCAAGCTCA 2256
 Qy 781 GlnValLeuProGlnValSerAlaGlyLysGln----- 791
 Db 2257 CAGTCTTGCTCAAGTATCAGCTGGAACAGAGGCTTCCACCTCGACTGCCACAG 2316
 Qy 791 ----- 791
 Db 2317 TACCCAGAGATTCAAATATTGCTCCTCTTCCACGAGGCTTCTGTGATCATCTT 2376
 Qy 791 ----- 791
 Db 2377 ACAGTCTTACCCTTCCCATGCCAGACAGAACTGCTACACCTGGGTACTTCCACA 2436
 Qy 791 ----- 791
 Db 2437 GTGTCAGACCTTATGTGGAATCAATCTTTTAACTTCTATGGTGTGTTAGAGAGAG 2496
 Qy 791 ----- 791
 Db 2497 GTTCAAGTGTCCAGCAGAGAGAGGATTAGACAAAGCCCCACTACATCTCCAGCAA 2556
 Qy 792 -----SerThrGlnGlyValSerGlnValAlaProAlaGlnProValAlaVal 807
 Db 2557 GCACTTTGAGAGATCTACAGGAGTCTCTCAGGTCTCTCTGAGAGCCAGTTGCACTA 2616
 Qy 808 AlaGlnProGlnAlaThrGlnProThrThrLeuAlaSerSerValAspSerAlaHisSer 827
 Db 2617 GCAAGCCCCAAGCTACCCAGCCAGCACTTGGCTTCTCTGTGACAGTGCACATTTCA 2676
 Qy 828 AspValAlaSerGlyMetSerAspGlyAsnGlnAsnValProSerSerSerGlyArgHis 847
 Db 2677 GATGTGCTTCAAGGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2736
 Qy 848 GluGlyArgThrThrLysArgHisTyrArgLysSerValAspSerArgSerArgHisGln 867
 Db 2737 GAAGAGAAACCTAAACAAACCGCATTTACCGAAATCTGTAGAGAGTGTGCGACATGAA 2796
 Qy 868 LysThrSerArgProLysLeuArgHLeuLeuValSerAsnLysGlyAspArgValVal 887
 Db 2797 AAAAATTACGCCCAAAATTAAGATTTTGAATGTTTCAATTAAGAGAGAGAGAGAG 2856
 Qy 888 GluCysGlnLeuGlnThrHisAsnArgLysMetValThrPheLysPheAspLeuAspGly 907
 Db 2857 GAATGTCAATTAAGACTCATTAATGAGAAATGTTTACATTCATTAATTTGACCTGATGAT 2916
 Qy 908 AspAsnProGlnGlnHLeuAlaThrHLeuThrValAsnAsnAspPheHLeuAlaHLeu 927
 Db 2917 GACAAACCCCGAGAGATGCAACCAATTAATGAGAAATGATTAATTTCTGCAATTAAG 2976
 Qy 928 ArgGlnSerPheValAspGlnValArgGlnHLeuGlyLysAlaAspGlnMetLeuSer 947
 Db 2977 AGAGAGTCTTGTGTGATCAAGTGCAGAAATTAATGAAAAAGCTGATGAATGCTCAGT 3036

QY 948 GluAspValSerValGluProGluGlyAspGlnGlyLeuGluSerLeuGlnGlyLysAsp 967
DB 3037 GAGGATGTCAGTGTGGAAACAGAGGGTGTATCAGGATTTGGAGTCTACAGAGAAAGAT 3096
QY 968 AspTyrGlyPheSerGlySerGlnLysLeuGlnGlyGluPheLysGlnProIleProAla 987
DB 3097 GACTATGGCTTTTCAGGTTCTCAGAAATGGAAAGGAGATTCAAAACAACTATTCTGCG 3156
QY 988 SerSerMetProGlnGlnIleGlyIlePheThrSerSerLeuThrGlnValAlaHisSer 1007
DB 3157 TCTTCATGCCACAGCAAAATAGCATTTCTACAGTTCTTTAACTCAAGTTTTCATTTCT 3216
QY 1008 AlaGlyArgArgPheIleValSerProValProGluSerArgLeuArgGluSerLysVal 1027
DB 3217 GCGGGAAGCGGCTTATAGTAGTCTGTCGCAAAAGCCATTACGGAATCAAAAGTT 3276
QY 1028 PheProSerGluIleIleThrAspThrValAlaAlaSerThrAlaGlnSerProGlyMetAsn 1047
DB 3277 TTCCCAAGTAAATACAGATACAGTTGCTGCTTACAGCTCAGAGCCCTGGAATGAAC 3336
QY 1048 LeuSerHisSerAlaSerSerLeuSerLeuGlnGlnAlaPheSerGluLeuArgArgAla 1067
DB 3337 TTGTCTCACTCGCATCATCTCTTAGTCTACACAGGCTTTTCTGAACCTTAGAGCTGCC 3396
QY 1068 GlnMetThrGlnGlyProAsnThrAlaProProAsnPheSerHisThrGlyProThrPhe 1087
DB 3397 CAATGACAGAAAGACCCAAVACAGACCTCCAAACTTAGTCATACAGAACCAACTTT 3456
QY 1088 ProValValProProPheLeuSerSerIleAlaGlyValProThrThrAlaAlaIleThr 1107
DB 3457 CCAAGTAGACCTCTCTTTAAAGTAGCATGCTGAGGTCACCAACACAGCAGCAGCCCA 3516
QY 1108 AlaProValProAlaThrSerSerProProAsnAspIleSerThrSerValIleGlnSer 1127
DB 3517 GCACAGTCCCTGCAACAGACAGCCCTCTATACATTTCCATCCATGTAATTCAGTCT 3576
QY 1128 GluValThrValProThrGlnGlnGlyIleAlaGlyValAlaIleThrSerThrGlyValVal 1147
DB 3577 GAGGTTCAGTGGCCCATGAGAGGGGATGCTGGAGTTGCCACAGACAGGTGTGTA 3636
QY 1148 ThrSerGlyGlyLeuProIleProProValSerGluSerProValLeuSerSerValVal 1167
DB 3637 ACTTCAGGTGGTCTCCCAATACACACTGTGTCTGATATCACAGTACCTTCCAGGGTGT 3696
QY 1168 SerSerIleThrIleProAlaValAlaValSerIleSerThrThrSerProSerLeuGlnVal 1187
DB 3697 TCAAGTATCAAAATCTCGAGTGTCTCAATATCTACTATACCCCGTCACTTCAAGTC 3756
QY 1188 ProThrSerThrSerGluIleValValSerSerThrAlaLeuTyrProSerValThrVal 1207
DB 3757 CCCACATCCACATCTGAGATCGTTGTTCTAGTACAGACATGTAATCTTCAGTAACGTT 3816
QY 1208 SerIleIleThrSerAlaSerAlaGlySerThrAlaThrProGlyProLysProProAla 1227
DB 3817 TCACCAACTTCAGCTCTGCGGGGGGAGTACTGTAATCCCAAGGTCTTAAAGCTTCCAGCT 3876
QY 1228 ValValSerGlnGlnAlaAlaGlySerThrThrValGlyAlaThrLeuThrSerValSer 1247
DB 3877 GTACTATCTACGAGGAGGAGGAGGAGCATCTGTGGAGGACCATTAACATTCAGTTTCT 3936
QY 1248 ThrThrThrSerPheProSerThrAlaSerGlnLeuSerIleGlnLeuSerSerThrThr 1267
DB 3937 ACCACCACTTCATTTCCCAAGACAGCTTTCACAGCTGTCATTTAGCTTAGAGAGACT 3996
QY 1268 SerThrProThrIleuAlaGluThrValValValSerAlaHisSerLeuAspLysThrSer 1287
DB 3997 TCTACTCTACTTATAGCTGAAACCGTGTAGTTAGCCACATCACTAAGATAGACATCT 4056
QY 1288 HisSerSerThrThrGlyLeuAlaPheSerLeuSerAlaProSerSerSerSerPro 1307
DB 4057 CATAGCAGTACAACTGATTTGGCTTTCTCCCTCTCTGACCACTTCCTCTTCTCTCCT 4116
QY 1308 GlyAlaGlyValSerSerTyrIleSerGlnProGlyGlyLeuHisProLeuValIlePro 1327
DB 4117 GGAGCAGAGATGTCATGTAATTTCTCAGCTGTGGGCTCCATCTTTGATCATTTCCA 4176
QY 1328 SerValIleAlaSerThrProIleLeuProGlnAlaAlaGlyProThrSerThrProLeu 1347
DB 4177 TCAGTGAATAGCTTCTACTCTCTATTTCTTCCCAAGACAGAGACCTTCTTCACTTCA 4236
QY 1348 LeuProGlnValProSerIleProProLeuValGlnProValAlaAsnValProAlaVal 1367
DB 4237 TTACCCCAAGAACCTAGATATCCACCTGTGTAGACGCTGTGTGCCAATGTGCTCTGTA 4296
QY 1368 GlnGlnThrIleuIleHisSerGlnProGlnProAlaLeuLeuProAsnGlnProHisThr 1387
DB 4297 CAGCAGACATTAATTCATATGTCAGCTTCAACAGCTTTGCTTCCACACAGCCCATCT 4356
QY 1388 HisCysProGluValAspSerAspThrGlnProLysAlaProGlyIleAspAspIleLys 1407
DB 4357 CATGTCTCTGAAGTGAATTTGATACACACCAAGCTCCTGGAAATGATGACATAAAG 4416
QY 1408 ThrLeuGlnGlyLysLeuArgSerLeuPheSerGluHisSerSerSerGlyAlaGlnHis 1427
DB 4417 ACTTAGAAGAAAGCTGCGGTCTCTGTTCAGTGAACACAGCTCATCTGAGCTCAGCAT 4476
QY 1428 AlaSerValSerLeuGluThrSerLeuValIleGlnSerThrValThrProGlyIlePro 1447
DB 4477 GCCTCTGCTCACTGAGACTCTCACTGATCATAGAGACACTGTACACAGGACATCCCA 4536
QY 1448 ThrThrAlaValAlaProSerLysLeuLeuThrSerThrThrSerThrCysLeuProPro 1467
DB 4537 ACTACTGCTGTGACCAACCAAGCAACTCTGACTTACCAACAAGTCTTGCTTACCAACA 4596
QY 1468 ThrAsnLeuProLeuGlyThrValAlaLeuProValThrProValValThrProGlyGln 1487
DB 4597 ACCAATTTACACATGAGAAACAGTGTCTTGCAGTTACACAGAGGTGCACACTGGGCA 4656
QY 1488 ValSerThrProValSerThrThrThrSerGlyValLysProGlyThrAlaProSerLys 1507
DB 4657 GTTTCATCCCAAGCAGACATCTCACTACATCAGAGATGAACCTGGAACTGCTCCCTCAG 4716
QY 1508 ProProLeuThrLysAlaProValLeuProValGlyThrGlnLeuProAlaGlyThrLeu 1527
DB 4717 CCACCTTAATAAGCTCCGCTCGCTCCAGTGGGTACTGAATCTTCCAGCAGGTACTCTA 4776
QY 1528 ProSerGlnGlnLeuProProPheProGlyProSerLeuThrGlnSerGlnGlnProLeu 1547
DB 4777 CCACGACAGAGCTGCCACCTTTCCAGAACCTTCTTAAACCACTCCACGAACTCTTA 4836
QY 1548 GluAspLeuAspAlaGlnLeuArgArgThrLeuSerProGlu**IleThrValThrSer 1567
DB 4837 GAGGATCTTATGCTCAATTAGAAGAAACCTTAGTCCAGATKATACAGTACTTCT 4896
QY 1568 AlaValGlyProValSerMetAlaAlaProThrAlaIleThrGlnAlaGlyThrGlnPro 1587
DB 4897 GCGGTGTGCTCTGTCATAGTCAAGGCGGCTCCAAACAGCAATCACAGAAACAGAGCT 4956
QY 1588 GlnLysGlyValSerGlnValLysGlnGlyProValLeuAlaThrSerSerGlyValGly 1607
DB 4957 CAGAGGGGTCTTCTCAGTCAAGTCAAGAGAGGCTGTCTTCACTCACTATTCAGAGGCTGT 5016
QY 1608 ValPheLysMetGlyArgPheGlnValSerValAlaAlaAspGlyAlaGlnLysGlnGly 1627
DB 5017 GTTTTAAAGTAGGAGCATTTTCAGGTTTCTGTTCGACGACAGGTCGCCGAAAGAGGCT 5076
QY 1628 LysAsnLysSerGluAspAlaLysSerValHisPheGlnSerSerThrSerGluSerSer 1647
DB 5077 AAAAATAGTCAAGATGCAAGATGCTGTTCAATTTCAATCCAGCACTCAGAGTCTCTCA 5136
QY 1648 ValLeuSerSerSerSerProGluSerThrLeuValLysProGluProAsnGlyIleThr 1667
DB 5137 GTGCTATCAAGTATGATGTCAGAGATCACTTGCTGAACCAAGCCGATGGCAATACCA 5196
QY 1668 IleProGlyLysSerSerAspValProGluSerAlaHisLysThrThrAlaSerGluAla 1687

Db 5197 ATCCCTGGTATCTCTTCAGATGTCAGAGAGTGCCACAAACTACTGCTCAGAGCA 5256
 Qy 1688 LysSerAspThrGlyGlnProThrLysValGlyArgPheGlnValThrThrAlaAsn 1707
 Db 5557 AAGTCACACACTGGGCGAGCTTACCAAGGTTGACGTTTCAAGGTGACAACTACGCAAC 5316
 Qy 1708 LysValGlyArgPheSerValSerLysThrGlnAspLysIleThrAspThrLysLysGlu 1727
 Db 5517 AAGGTGGTGCTTCTCTGTATCAAAAAGTACAGACAAAGATCACTGACACAAAGAAAGAA 5376
 Qy 1728 GlyProValAlaSerProProPheMetCAspLeuGlnAlaValLeuProAlaValIle 1747
 Db 5377 GACACAGTGGATCTCTCTCTTATGATTTGGAAACAGCTGTTCTTCTGCTGTATA 5436
 Qy 1748 ProLysLysGlyLysProGlnLeuSerGluProSerHisLeuAsnGlyProSerSerAsp 1767
 Db 5437 CCAAGAGAAAGAGAGCTGAACTGTACAGCTTCACTTAATAGGCGCTCTTCTGAC 5496
 Qy 1768 ProGlnAlaAlaPheLeuSerArgAspValAspAspGlySerGlySerProHisSerPro 1787
 Db 5497 CCGGAGGCGCTTTTAAAGTAGGAGTGTGATGATGTTCCGGTAGTCCACACTCGGCC 5556
 Qy 1788 HisGlnLeuSerSerLysSerLeuProSerGlnAsnLeuSerGlnSerLeuSerAsnSer 1807
 Db 5557 CATCAGCTGAGCTCAAGAGCTTCTTCAAGCAATCTTAAGTCAAGGCTTAAATTTCA 5616
 Qy 1808 PheAsnSerSerLysMetSerSerSerAspAsnGlnSerAspIleGlnAspGluAspLeuLys 1827
 Db 5617 TTTAACTCTCTTACATAGATAGGACATAGTCAGATACGATACGAAAGTAAAGACTTAAAG 5676
 Qy 1828 LeuGlnLeuArgArgLeuArgAspLysHisLeuLysGlnIleGlnAspLeuGlnSerArg 1847
 Db 5677 TTAAGAGCTGGACGACATACAGATTAACATCTCAAGAGATTCAGGACTTCAGAGTGC 5736
 Qy 1848 GlnLysHisGlnIleGlnSerLeuThrLysLeuGlyLysValProProAlaValIle 1867
 Db 5737 CAGAAAGCATGAATGTAATCTTTGATACCAAACTGGGCAAGGCGCCCTGCTGTTATT 5796
 Qy 1868 IleProProAlaAlaProLeuSerGlyArgArgArgProThrLysSerLysGlySer 1887
 Db 5797 ATTCCTCCAGCTGCTCCCTTCAGGAGAAAGACACACCACTTAAGCAAAAGGACGC 5856
 Qy 1888 LysSerSerArgSerSerSerLysGlyAsnLysSerProGlnLeuSerGlyAsnLeuSer 1907
 Db 5857 AAATCTAATCGAAGAGATTCTTGGGGAATAAAGCCCCAGCTTTCAGTAACCTGCT 5916
 Qy 1908 GlyGlnSerAlaAlaSerValLeuHisProGlnGlnThrLeuHisProProGlyAsnIle 1927
 Db 5917 GGTCAAGATGAGCTTCAAGTCTTGACCCCAAGAGACCTCCACCTCTGCGCAACATC 5976
 Qy 1928 ProGlnSerGlyGlnAsnGlnLeuGlnProLeuLysProSerProSerSerAspAsn 1947
 Db 5977 CCAGAGTCCGGGCAAAATGAGTGTACAGCCCTTAAGCCATCTCCCTCCAGTGACAAAC 6036
 Qy 1948 LeuThrSerAlaPheThrSerAspGlyAlaIleSerValProSerLeuSerAlaProGly 1967
 Db 6037 CTCTAATTCAGCTTACCAAGATGTGTCCATTTCAGTACCAAGCTTTCGCTCAGAT 6096
 Qy 1968 GlnGlyThrSerSerThrAsnThrValGlyAlaThrValAsnSerGlnAlaGlnAla 1987
 Db 6097 CAAGGAACACGACGACAAACCTGTTGGGCAACAGTGAAACAGCAAGCCGCCCAACT 6156
 Qy 1988 GlnProProAlaMetThrSerSerArgLysGlyThrPheThrAspAspLeuHisLysLeu 2007
 Db 6157 CAGCTCTCTGCATACGTCACGACGAGAAAGGCAATTCACAGATGACTGCAAGTTG 6216
 Qy 2008 ValAspAsnThrAlaArgAspAlaMetAsnLeuSerGlyArgArgGlySerLysGlyHis 2027
 Db 6217 GTAGACATTTGGGCGCCGAGATGCCATGAATCTCTACAGAGAGAGAGAAAGGAGCAC 6276
 Qy 2028 MetAsnThrGlyGlyProGlyMetAlaArgLysPheSerAlaProGlyGlnLeuGlyIle 2047
 Db 6277 ATGAATTTAGAGGCGCTGGAAATGCAAGAAATTTCTTGACCTGGGCAACTGTGCAATC 6336

Qy 2048 SerMetThrSerAsnLeuGlyGlySerAlaProIleSerAlaAlaSerAlaThrSerLeu 2067
 Db 6337 TCCATGACCTCGAAGCTGGGTGCTGCGCCCATCTCTGACATCAGATCAGTACTCTCTA 6396
 Qy 2068 GlyHisPheThrLysSerMetCysProProGlnGlnThrGlyPheProAlaThrProPhe 2087
 Db 6397 GGTCACTTCAACAGTCTATGTCGCCCCCAACACAGATGATGCTTCCAGCTACCCCATTT 6456
 Qy 2088 GlyAlaGlnThrPheSerGlyThrGlyGlyProAlaProGlnProLeuGlyGlnPheGlnPro 2107
 Db 6457 GGGCTCAATGAGTGGGACGGGTGGCCGACACACAGCACTTGGCCAGTTCCAACT 6516
 Qy 2108 ValGlyThrAlaSerLeuGlnAsnPheAsnIleSerAsnLeuGlnLysSerIleSerAsn 2127
 Db 6517 GTGGAACTGCTCTCTTGCAGAAATTTCAATCAGCAATTTGGAGAAATCATCAGCAAC 6576
 Qy 2128 ProProGlySerAsnLeuArgThrThr 2136
 Db 6577 CCCCAGCTCCAACTGCGGACCACT 6603

RESULT 9

ACH03772 standard; DNA; 7065 BP.

ACH03772;

26-SEP-2003 (first entry)

DNA encoding novel human protein kinase #9.

Human; kinase; ds; gene; cosmetic application; nutraceutical application.

Homo sapiens.

US6541252-B1.

01-APR-2003.

14-MAY-2001; 2001US-00854856.

19-MAY-2000; 2000US-0206015P.

(LEXI-) LEXICON GENETICS INC.

Walke DW, Hillbun E, Donoho G, Turner CA;

WPI; 2003-575927/54.

P-PSDB; ABO44389.

New nucleic acid encoding novel human proteins, useful in cosmetic and nutraceutical applications.

Disclosure; Page; 11pp; English.

The invention relates to a new isolated nucleic acid encoding a novel human protein kinase. The nucleic acid is useful in cosmetic and nutraceutical applications. The present sequence represents DNA encoding a novel human protein kinase. Note: The sequence data for this patent did not from part of the printed specification but was obtained in electronic format directly from USPTO at seqdata.uspto.gov/sequence.html?docid=6541252B1

Sequence 7065 BP; 1911 A; 1940 C; 1601 G; 1610 T; 0 U; 3 Other;

Alignment Scores:

Pred. No.:	0	Length:	7065
Score:	10513.00	Matches:	2108
Percent Similarity:	88.50%	Conservative:	0
Best Local Similarity:	88.50%	Mismatches:	0
Query Match:	97.23%	Indels:	274
DB:	8	Gaps:	2

US-10-010-720-14 (1-2136) x ACH03772 (1-7065)

```

QY 1 MetSerGlyValAlaGluValSerThrProGlySerLeuPheLeuSerPro 20
Db 1 ATGTCTGGCGGGCGCGCAGAGAGAGACACACTCCCGGTTCTCTCTCGCGG 60
QY 21 ProAlaProAlaProIysAsnGlySerSerSerSerSerValGlyValIleuGly 40
Db 61 CGCGCTCTCGCCCCCAAGAAATGGCTCCAGCTCCGATTTCTCCGTGGGGAGAAACTGGGA 120
QY 41 AlaAlaAlaAlaAspAlaValThrGlyValThrGlyValThrGlyValThrGlyValThr 60
Db 121 GCGCGGGCGCGCGAGCTGTGACCGCGAGGACCGAGAGTACAGCGCGCGCGCACACT 180
QY 61 MetAspLysAspSerArgGlyValAlaAlaThrThrThrThrThrGluHisArgPhePhe 80
Db 181 ATGGAGCAAGACAGCCGTGGGGCGCGCGCACCTTACACCACTGAGACCGCTTCTTC 240
QY 81 ArgArgSerValIleCysAspSerAsnAlaThrAlaLeuGluLeuProGlyLeuProLeu 100
Db 241 CGCCGAGCGCTCATCTGCGACTCCCAATGCACTGCACTGAGCTTCCGGGCTTCTCTT 300
QY 101 SerLeuProGluProSerIleProAlaAlaValProGluSerAlaProProGluProHis 120
Db 301 TCCCTGCCCCAGCCCGACGATCCCGGGCTGTCCGAGAGTGTCCACCGAGCGCCAC 360
QY 121 ArgGluGluThrValThrAlaThrSerGluValAlaGluGluProProAlaAla 140
Db 361 CGGAGAGAGACCGTGCACCGCCACCGCCACTTCCAGGTAAGCCCGACCGCTCCACCGCT 420
QY 141 AlaAlaProGlyGluGluAlaValAlaGlyProAlaProSerThrValProSerSerThr 160
Db 421 GCCCGCCCTGGGGAGAACAGCGCGCGCGCGCTGCGCTCCGACTGTCCCGACGAGTACC 480
QY 161 SerLysAspAspProValSerGluProSerLeuValGlySerLysGluGluProProPro 180
Db 481 AGCAAAAGCCCGCCAGAGTCCAGCTTAGCTTGGGGAGCAAAAGAGAGACCGCGCGG 540
QY 181 AlaArgSerGlySerGlyGlySerAlaLysGluProGluGluGluArgSerGluGlu 200
Db 541 GCGAGAGTGGCAGCGCGCGCGCGCGCCCAAGAGCCACAGAGAGAGAGAGAGAGAGAG 600
QY 201 GluAspAspIleGluGluLeuGluThrLysAlaValGlyMetSerAsnAspGlyArgPhe 220
Db 601 CAGGATGATATCGAAGAGCTGAGACCAAGCGCGTGGGAATGTCTAACAGATGGCGCTT 660
QY 221 LeuLysPheAspIleGluIleGlyArgGlySerPheLysThrValTyrLysGlyLeuAsp 240
Db 661 CTCAAGTTTGACATCGAAATCGGAGAGGCTCTTTAAGACGGTCTCAAAAGGCTCTGAC 720
QY 241 ThrGluThrThrValGluValAlaTyrCysGluLeuGluAspArgLysLeuThrLysSer 260
Db 721 ACTGAAACCAACCGTGAAGTGGCTGTGTGCAATCGAGATCGAAATTTAACAAAGTCT 780
QY 261 GluArgGluArgPheLysGluGluAlaGluMetLeuLysGlyLeuGluHisProAsnIle 280
Db 781 GAAGAGGAGAGATTAAAGAAAGAGCTGAATGTTAAAGGTCTTCAGCATCCCAATATT 840
QY 281 ValArgPheTyrAspSerTyrGluSerThrValLysGlyLysLysCysIleValLeuVal 300
Db 841 GTTAGATTATATGATCTCTGGGAATCCAGATAAAAGAAAGAGTGCATTGTTTGGTG 900
QY 301 ThrGluLeuMetThrSerGlyThrLeuLysThrTyrLeuLysArgPheLysValMetLys 320
Db 901 ACTGAACTTATGACGTCTGGAAACCTTAAACGATCTGAAAAAGTTTAAAGTATAG 960
QY 321 IleLysValLeuArgSerTyrCysArgGluIleLysGlyLeuGluPheLeuHisThr 340
Db 961 ATCAAGTTCTAAGAGAGCTGTGCGCTCAATCTTAAAGTCTTCAAGTTCTTCAACT 1020
QY 341 ArgThrProProIleIleHisArgAspLeuLysCysAspAsnIlePheIleThrGlyPro 360
Db 1021 CGAACTCAACCTATCATTCACCGGATCTTAAATGTGACAAACATCTTTATCACCGGCTT 1080

```

```
QY 721 GlnProSerSerSerSerLeuThrGlyValSerSerSerGlnProIleGlnHisProGln 740
Db 2139 ----- 2139
QY 741 GlnGlnGlnGlyIleGlnGlnThrAlaProProGlnGlnThrValGlnIleYrSerLeuSer 760
Db 2140 ---CAGCAGGAGAAATACAGCAGACAGACAGCCCTCCTCAACAGACAGACAGATATTCACCTTCA 2196
QY 761 GlnThrSerThrSerSerSerGlnAlaThrThrAlaGlnProValSerGlnProGlnAlaPro 780
Db 2197 CAGACATCACTCACTCACTGAGGAGCACTATCTCAACAGACAGATGAGTCAAGCTCAAGCTCCA 2256
QY 781 GlnValLeuProGlnValSerAlaGlyLysGln----- 791
Db 2257 CAAGCTTGCTCAAGATACAGCTGGAAGAAACAGCTTCAGTTTCCAGCAGTACCAACT 2316
QY 791 ----- 791
Db 2317 ATCCAGGCGAAGCTCAGATCCAGTTGCGACAGAACCTCGGTTGTTCCAGTCCACTCT 2376
QY 791 ----- 791
Db 2377 GGTGCTCATTTCTCTCCAGTGGAGACAGCCGCTCCCTACTCCCTGCTCCTCAGTACCT 2436
QY 791 ----- 791
Db 2437 GTCTCTCAGATTCCCATATCAACTCTCATGTGTCTACGGCTCAGACAGAGTTTCTCATCC 2496
QY 791 ----- 791
Db 2497 CTTCATCATCAATGCGAGCTGGCATTTACTCAGCTCTGTCTCAGGTGGCTTCACTGCT 2556
QY 791 ----- 791
Db 2557 ACAACAGCTGGAGTCCGAGGAGTATCAACTGTGTCTCAGTCAAGCTTCAACCTTCTG 2616
QY 791 ----- 791
Db 2617 CAGCTGTGACTCAGCTGCCAAGTCAAGTTCAACCAAGCTCTTCAACACAGAGTTGAG 2676
QY 791 ----- 791
Db 2677 TCCATGGGAATACAGCTAACCTTGGACAGCTGCTGAGTTCCACTTCTCTGAGAT 2736
QY 791 ----- 791
Db 2737 GTTCTGTACAGGGCTTCCCACTGACTGCCACCAAGTACCAAGAGATTCAAATATT 2796
QY 791 ----- 791
Db 2797 GCTCCCTCTTCCAGTGGCTTCTGTTCGATCCATTCACAGTCTTAMCCCTCCCATG 2856
QY 791 ----- 791
Db 2857 CCGACAGAGTACTGGCTACACCTGAGTACTTTCCACAGTGTGACGCTTATGTGAA 2916
QY 791 ----- 791
Db 2917 TCAAACTTTTAGTTCTTAGTGGGTGTGTAGAGAGACAGTTCAAGTCTCCAGCCAGAG 2976
QY 792 ----- 792
Db 2977 GGGAGTTTAGCAAGAGCCCACTACATCTCCAGCAAGAGAGTTTGGAGAGTACTGAG 3036
QY 795 GlyValSerGlnValAlaProAlaGluProValAlaValAlaGlnProGlnAlaThrGln 814
Db 3037 GGAGCTCTCTCAGGTTGCTCTCCAGAGCAGATTGACAGTACAGACCCCAACCTACCCAG 3096
QY 815 ProThrThrLeuAlaSerSerValaSerAlaHisSerAspValAlaSerGlyMetSer 834
Db 3097 CCGACCACTTGGCTTCTCTGTAGACAGTGCACATTCAGATGTGCTTCAGAGTATGATC 3156
QY 835 AspGlyAsnGluAsnValProSerSerSerGlyArgHisGluGlyArgThrThrIleArg 854
Db 3157 GATGGCAATGAGAACGCCATCTTCAGAGGAGGAGCATAGAGAAAGAACTACAAACCG 3216
QY 855 HisThrArgLysSerValaArgSerArgSerArgHisGluIleYrThrSerArgProLysLeu 874
Db 3217 CATTAACGAAATCTGTAGAGAGTCTCTCGACATGAAAAAACTTACGCCCAAAATTA 3276
QY 875 ArgIleLeuAsnValSerAsnLysGlyAspArgValaGluCysGlnLeuGluThrHis 894
Db 3277 AGAATTTTGAATGTTTCAATTAAGAGAGCCAGATAGATGATGATTCATTAATGAGATCAT 3336
QY 895 AsnArgLysMetValThrPheLysPheAspLeuAspGlyAspAsnProGluGluIleAla 914
Db 3337 AATAGGAAAAATGGTTACATTCAAATTTGACCTATGATGTACAAACCCGAGAGATGCA 3396
QY 915 ThrIleMetValaAsnAsnAspPheIleLeuAlaIleGluArgGluSerPheValaAspGln 934
Db 3397 ACAATTTATGGTAAACATGACTTTTATTTAGCAATAGAGAGAGAGTGTGTGAGATCA 3456
QY 935 ValArgGluIleIleGluLysAlaAspGluMetLeuSerGluAspValSerValaGluPro 954
Db 3457 GTGGAGAAATTAATGAAAAAGCTGATGAATGCTCAGTGAATGTCAGTGTGGAACCA 3516
QY 955 GluGlyAspGlnGlyLeuGluSerLeuGlnGlyLysAspAspIleYrGlyPheSerGlySer 974
Db 3517 GAGGTGATCAGGAGATTGAGAGCTTACAGAGAAAGATGACTATGGCTTTCAAGTCT 3576
QY 975 GlnLysLeuGluGluGluPheLysGlnProIleProAlaSerSerMetProGlnIle 994
Db 3577 CAGAAATTTGAGAGAGAGTTCAAACCAACAAATCTGTGCTTCTCAATGCCACAGCAATA 3636
QY 995 GlyIleProThrSerSerLeuThrGlnValaHisSerAlaGlyArgArgPheIleVal 1014
Db 3637 GGCAATCTCAACATTCCTTTAATCAAGTTGTTCACTCTCGGAGAGCGGTTATATG 3696
QY 1015 SerProValProGluSerArgLeuArgGluSerLysValPheProSerGluIleThrAsp 1034
Db 3697 AGTCTGTGCCAGAAAGCCGATTACGAGAAATCAAAAGTTTCCACAGGAAATTAACAT 3756
QY 1035 ThrValAlaAlaSerThrAlaGlnSerProGlyMetLeuSerHisSerAlaSerSer 1054
Db 3757 ACAGTGTGCTCTCAAGCTCAAGGCTTGAAGTGAATGAACTGTCTCACTCTGCATATCC 3816
QY 1055 LeuSerLeuGlnGlnAlaPheSerGluLeuArgArgAlaGlnMetThrGluLysProAsn 1074
Db 3817 CTTAGTCTACAGACGCTTTTCTGAATTTGACGTGCCAAATGACAGAGAACCAAY 3876
QY 1075 ThrAlaProProAsnPheSerHisThrGlyProThrPheProValaProProPheLeu 1094
Db 3877 ACAGACCTCCAAACTTATGATACATACAGAGCAACATTTCCAGTATGACCTCTTCTTA 3936
QY 1095 SerSerIleAlaGlyValProThrThrAlaAlaThrAlaProValProAlaThrSer 1114
Db 3937 AGTACATTTGCTGAGTCCCAACACAGACAGACAGCCACAGACAGTCCCTGCAACAGC 3996
QY 1115 SerProProAsnAspIleSerThrSerValIleGlnSerGluValThrValProThrGlu 1134
Db 3997 AGCCTCTTAATGACATTTCCACATCAGTATTCAGTCTGAGGTATCAGTCCCACTGAA 4056
QY 1135 GluGlyIleAlaGlyValAlaThrSerThrGlyValAlaThrSerGlyLysProIle 1154
Db 4057 GAGGGAGATTTGCTGAGTGCACACAGACAGGTGTGTTACTTCAGGTGCTTCCCAATA 4116
QY 1155 ProProValSerGluSerProValLeuSerSerValaSerSerIleThrIleProAla 1174
Db 4117 CCACCTGTGTCTGAATCCACGATCTTCTACAGGATTTCAAGTATCAACATTAATCTGGA 4176
QY 1175 ValValSerIleSerThrThrSerProSerLeuGlnValProThrSerThrSerGluIle 1194
Db 4177 GTTGCTCAATATCTACTACATCCCGTCACTTCAAGTCCCACTCACATCTAGATC 4236
QY 1195 ValValSerSerThrAlaLeuIleYrProSerValThrValSerAlaThrSerAlaSerAla 1214
```


Db 4237 GTTGTTCAGTACAGCACTGATTCCTTCAGTAACAGTTTCAGCACTTCAGCTCTGCA 4296
QY 1215 GLYISerThrAlaThrProGlyProLysProProAlaValSerGlnGlnAla 1234
Db 4297 GGGGGCAGTACGTCTACCCCAAGTCTTAAGCTCCAGCTGTAGTATCTAGAGAGAGCA 4356
QY 1235 GLYSerThrValGlyAlaThrLeuThrSerValSerThrThrThrSerPheProser 1254
Db 4357 GGCAGACACTAGTGGAGCCACATTAACATGATTTCTACACCACTTCATCCCAAGC 4416
QY 1255 ThrAlaSerGlnLeuSerLeuGlnLeuSerSerSerThrThrProThrLeuAlaGlu 1274
Db 4417 ACAGCTTACAGCTGTCCATTCAGCTTCAGACAGACTTCATTCCTCACTTAGCTGAA 4476
QY 1275 ThrValValSerAlaHisSerLeuAspLysThrSerHisSerSerThrThrGlyLeu 1294
Db 4477 ACCGGTAGTAGTACGGCACACTCATTAGATTAAGATTCATTAAGCACTGAACTTG 4536
QY 1295 AlaPheSerLeuSerAlaProSerSerSerSerProGlyAlaGlyValSerSerTyr 1314
Db 4537 GCTTCTCCCTCTGCAACATCTTCCTCTCTCTCCGAGAGAGGTGCTAGTAT 4596
QY 1315 IleSerGlnProGlyGlyLeuHisProLeuValIleProSerValIleAlaSerThrPro 1334
Db 4597 ATTCTCAGCTGTGGGCTGATCTTGTGTCATTCATCACTAGATAGCTTCATCTCT 4656
QY 1335 IleLeuProGlnAlaAlaGlyProThrSerThrProLeuLeuProGlyValProSerIle 1354
Db 4657 ATTCTCCCAAGAGCAGAGCACTTCTACACTTTATTCCTCCCAAGTACTTCATCTCT 4716
QY 1355 ProProLeuValGlnProValAlaAsnValProAlaValGlnGlnThrLeuIleHisSer 1374
Db 4717 CCACCTTGGTACAGCTGTGCTGCAATGTGCTGCTGTAAGACAGACACTAATTCATGT 4776
QY 1375 GlnProGlnProAlaLeuLeuProAsnGlnProHisThrHisGlySerProGlyValAspSer 1394
Db 4777 CAGCTTCAACAGCTTGTCTCTCCCAACCAAGCCCACTCATGTGCTCGAAGTAAATTC 4836
QY 1395 AspThrGlnProLysAlaProGlyIleAspAspIleLysThrLeuGlnGlyLysLeuArg 1414
Db 4837 GATACACAAACCAAGCTCTGGAATGATGACATTAAGACTCTGAAAGAAAGCTGCGG 4896
QY 1415 SerLeuPheSerGlnHisSerSerSerSerGlyAlaGlnHisAlaSerValSerLeuGlnThr 1434
Db 4897 TCTGTGTCAAGTGAACAGCTCATCTGAGCTCAGCATGCTGTCTCACTGGAAGCC 4956
QY 1435 SerLeuValIleGlnSerThrValThrProGlyIleProThrThrAlaValAlaProSer 1454
Db 4957 TCACCTAGTACATAGAGAGCACTGTACACAGCAGCATCCCACTACTGCTGTGCAACAGC 5016
QY 1455 LysLeuLeuThrSerThrThrSerThrCysLeuProProThrAsnLeuProLeuGlyThr 1474
Db 5017 AAACCTCCGACTTCTTACCAAGATACCTTACTTACCAACCAATTTCACCTAGGAACA 5076
QY 1475 ValAlaLeuProValThrProValValThrProGlyGlnValSerThrProValSerThr 1494
Db 5077 GTTGTCTTGGCAGTTACACAGCTGTACACCTGGGCAAGTTTCTACCCCACTGAGCACT 5136
QY 1495 ThrThrSerGlyValLysProGlyThrAlaProSerLysProProLeuThrLysAlaPro 1514
Db 5137 ACTACATCAGAGTGAAGAACTGTGAAGCTCTCCCAAGCACTCTTAACCTTAAGGCTCGG 5196
QY 1515 ValLeuProValGlyThrGlnLeuProAlaGlyThrLeuProSerGlnGlnLeuProPro 1534
Db 5197 GTGTGTCCAGTGGTACTGAATCTCCAGCAAGTACTTACCCAGCGAGAGCTGCCACT 5256
QY 1535 PheProGlyProSerLeuThrGlnSerGlnGlnProLeuGlnAspLeuAspAlaGlnLeu 1554
Db 5257 TTTCACAGACCTTCTCTAACCCAGTCCCAAGCACTCTAGAGAGATCTTGATGTCTCAATTG 5316
QY 1555 ArgAlaThrLeuSerProGluLysIleThrValThrSerAlaValGlyProValSerMet 1574
Db 5317 AGAAGAACCTTAGTCCAGATATATCAAGTACTTCTGTGGTGTCTCTGTCTCATG 5376

QY 1575 AlaAlaProThrAlaAlaIleThrGlnAlaGlyThrGlnProGlnLysGlyValSerGlnVal 1594
Db 5377 GCGGCTCCAAACAGCATATCAAGAACAGAGAACAGCCCTCAGAAAGGTGTCTTCTCAAGTC 5436
QY 1595 LysGlnGlyProValLeuAlaThrSerSerGlyAlaGlyValPheLysSerGlyArgPhe 1614
Db 5437 AAGAAAGCCCTGTCTTACCACTAGTTCAGAGAGCTGGTGTTTTAAAGATGGAGCAATT 5496
QY 1615 GlnValSerValAlaAlaAspGlyAlaGlnLysGlnGlyLysAsnLysSerGlnAspAla 1634
Db 5497 CAGTTCTGTGTGACAGACAGCGTCCCAAGAAAGAGGGTAAATTAAGTCANAAGATGA 5556
QY 1635 LysSerValHisPheGlnSerSerThrSerGlnSerSerValLeuSerSerSerPro 1654
Db 5557 AAGTCTGTTCATTTGAATTCAGACACCTCAGAGTCTCAGTCTATCAATGATGTGCTCA 5616
QY 1655 GluSerThrLeuValLysProGlnProAsnGlyIleThrIleProGlyIleSerSerAsp 1674
Db 5617 GAGAGTACTTGTGTGAACCAAGCCGAATGGCATACATCTCTGTATCTTCTCAGAT 5676
QY 1675 ValProGlnSerAlaHisLysThrThrAlaSerGlnAlaLysSerAspThrGlyGlnPro 1694
Db 5677 GTGCCAGAGAGTGCCCAAAACTACTGCTCAGAGGCAAGTCAAGCTGGCAGCT 5736
QY 1695 ThrLysValGlyArgPheGlnValThrThrThrAlaAsnLysValGlyArgPheSerVal 1714
Db 5737 ACCAAGGTGAGCGTTTCAGGTGACCACTACAGCAACCAAGTGGTGTCTCTCTCTA 5796
QY 1715 SerLysThrGlnAspLysIleThrAspThrLysLysGlnGlyProValAlaSerProPro 1734
Db 5797 TCAAAACGTAGAGCAAGATACCTAGTACCAAGAAAGAAAGACAGTGGCATCTCTCTCT 5856
QY 1735 PheLeuAspLeuGlnGlnAlaValLeuProAlaValIleProLysLysGlnLysProGlu 1754
Db 5857 TTATGATTTGGAAAGAGCTGTCTCTCTCTGATACCAAGAAAGAAAGAGCTGAA 5916
QY 1755 LeuSerGlnProSerHisLysAsnGlyProSerSerAspProGlnAlaAlaPheLeuSer 1774
Db 5917 CTGTCAAGCTTTCACATCTTAATGGGCGCTTCTTGACCCGAGGCGCTTTTATAGT 5976
QY 1775 ArgAspValAspAspGlySerGlySerProHisSerProHisGlnLeuSerSerLysSer 1794
Db 5977 AGGATGTGATGTATGGTTCGGTAGTCCACACTGGCCCATCAGCTCAGCTCAAAAGCC 6036
QY 1795 LeuProSerGlnAsnLeuSerGlnSerLeuSerAsnSerPheAsnSerSerTyrMetSer 1814
Db 6037 CTTCCTAGCCAGATCTTAAGTCAAAAGCTTAGTAATTCATTAACTCTCTTACATGAGT 6096
QY 1815 SerAspAsnGlnSerAspIleGlnAspGlnAspLeuLysLeuGlnLeuArgArgLeuArg 1834
Db 6097 AGCCACATAGTATGATATGATGAAAGATGAAGACTTAAAGTTAAGGCGCAGCACTGCA 6156
QY 1835 AspLysHisLeuLysGlnIleGlnAspLeuGlnSerArgGlnLysHisGlnIleGlnSer 1854
Db 6157 GATTAACATCTCAAAAGAGATTTCAGGACTTCGAGAGTCCAGAAAGCATGAATTAATCT 6216
QY 1855 LeuThrThrLysLeuGlyLysValProProAlaValIleIleProProAlaAlaProLeu 1874
Db 6217 TTGTATACCAAACTGGGCAAGGTGCCCTGTGTATTTATTCCTCCCACTGCTCCCTT 6276
QY 1875 SerGlyArgArgArgArgProThrThrLysSerGlySerLysSerSerArgSerSerSer 1894
Db 6277 TCAGGGAAGAGAGAGCAACCCATTAAGAAAGCAAGGCAAAATCTATTCAGAGCAATTC 6336
QY 1895 LeuGlyAsnLysSerProGlnLeuSerGlyAsnLeuSerGlyGlnSerAlaAlaSerVal 1914
Db 6337 TTGGGGAATAAAGCCCCCAAGCTTTCAGTATCTGTCTGTCTCAGAGTGAAGCTTCAGTC 6396
QY 1915 LeuHisProGlnGlnThrLeuHisProProGlyAsnIleProGlnSerGlyGlnAsnGln 1934
Db 6397 TTGCACCCCAAGAGACCTCCACCTCTGAGCAACATCCCAAGTCCGGGAGAAATCAG 6456

QY 301 ThrGluMetThrSerGlyThrLeuThrTyLeuLeuArgPheLeuValMetLys 320
Db 721 ACTGAACCTTATGACGCTCTGGAACACTTAAAGTATCTGAAAGGTTTAAAGTGAAG 780
QY 321 IleValLeuLeuArgSerThrPyrArgGlnIleLeuLysGlyLeuGlnPheLeuHisThr 340
Db 781 ATCAAGATCTTAAGAGAGCTGGTGCCTGAGATCCTTAAAGGTCTTCAGTTTCTTCACT 840
QY 341 ArgThrProProIleIleHisArgAspLeuLysCysAspAsnIlePheIleThrGlyPro 360
Db 841 CGAATCTCACCTTATCATTCATCCGCGATCTTAAATGTGACACATCTTATACCCGCGCT 900
QY 361 ThrGlySerValLysIleGlyAspLeuGlyLeuValaThrLeuLysArgAlaSerPheAla 380
Db 901 ACTGCTCAGTCACAGATGAGACCTCGGTCTGGACACCTGAAACCGGGCTTCTTTGCC 960
QY 381 LysSerValIleGlyThrProGlnPheMetAlaProGlnMetTyrGlnGlnLysTyrAsp 400
Db 961 AAGAGTGTATGATGATGATCCCGAGATTCATGGCCCTGTGATGTATGAGAGAAATATGAT 1020
QY 401 GluSerValAspValTyrAlaPheGlyMetCysMetLeuGlnMetAlaThrSerGlnTyr 420
Db 1021 GAATCCGTTGACGTTTATGCTTTGGATGTGCACTGCTTGAATGGCTTACATCTGAATAT 1080
QY 421 ProTyrSerGlnCysGlnAsnAlaIleGlnIleTyrArgValThrSerGlyValLys 440
Db 1081 CCTTACTGAGGTGCGAAATGCTGCACAGATCTACCGTCCGGTGAACAGTGGGTGAAG 1140
QY 441 ProAlaSerPheAspLeuValAlaIleProGlnValLysGlnIleIleGlnGlyCysIle 460
Db 1141 CCAGCCAGTTTGGACAAAGTACATCTTGAAGAGAAATTTATGAAAGATGACATA 1200
QY 461 ArgGlnAsnLysAspGluArgTyrSerIleLysAspLeuLeuAsnHisAlaPhePheGln 480
Db 1201 CGACAAACAAAGATGAAAGATATTCCTCAAGACCTTTGAACCATGCTTCTTCCA 1260
QY 481 GlnGlnThrGlyValArgValGlnLeuAlaGlnLysAspGlyLysIleAlaIle 500
Db 1261 GAGAAACAGAGATGACGGGTGAATTAGCAGAAAGATGATGAGAAATAATGACATA 1320
QY 501 LysLeuThrPheLeuArgIleGluAspIleLysLeuLysGlyTyrTyrLysAspAsnGln 520
Db 1321 AATTTATGGCTACGATATTGAAGATTTAAGAAATTTAAAGGAGAAATACAAAGATTAAGAA 1380
QY 521 AlaIleGlnPheSerPheAspLeuGlnArgAspValProGlnAspValAlaGlnGlnMet 540
Db 1381 GCTATGATGATTTCTTTTGTGATTAGAGAGAGATCTCCAGAAAGATGTTGCACAAAGAAATG 1440
QY 541 ValGlnSerGlyTyrValCysGlnGlyAspHisLysThrMetAlaLysAlaIleLysAsp 560
Db 1441 GTAGAGCTGGGTATGCTGTGAGAGTGATCACAAGACCATGCTTAAAGCTATCAAAAGAC 1500
QY 561 ArgValSerLeuIleLysArgLysArgGlnGlnArgGlnLeuValArgGlnGlnGlnGln 580
Db 1501 AGAGTATCATTTAATTAAGAGAAACGAGAGCGGAGTTGTGACGAGAGACCAAGAA 1560
QY 581 LysLysLysGlnGlnGlnGlnSerSerLeuLysGlnGlnValGlnGlnSerSerAlaSerGln 600
Db 1561 AAAAAAAGCAGAGAGAGAGAGAGCTCAAAACAGAGGTAGAACATCAAGTCTTCCAG 1620
QY 601 ThrGlyIleLysGlnLeuProSerAlaSerThrGlyIleProThrAlaSerThrThrSer 620
Db 1621 ACGAGATCAAGACGCTCCCTTCTGCTAGCACCGGCTATCTACTGCTTCACTTCA 1680
QY 621 AlaSerValSerThrGlnValGlnProGlnGlnProGlnAlaAspGlnHisGlnGlnLeu 640
Db 1681 GCTTCACTTCTACACAAAGTGAACCTGAGAGACCTGAGGAGATCAACATCAACAACTA 1740
QY 641 GlnTyrGlnGlnProSerIleSerValLeuSerAspGlyThrValAspSerGlyGlnGly 660
Db 1741 CAGTACCAAGAACCCAGATATATCTGTATCTGATGGACGTTTACAGATGCTGACAGGA 1800
QY 661 SerSerValPheThrGlnSerArgValSerSerGlnGlnThrValSerTyrGlySerGln 680
Db 1801 TCCTCTGCTTCAAGAAATCTGAGTGAAGCAACCAAGACGTTTCATATGATGCCAA 1860
QY 681 HisGlnGlnAlaHisSerThrGlyThrValProGlnHisIleProSerThrValGlnAla 700
Db 1861 CATGAACAGGACATCTTACAGGACAGATGCCAGGAGCATATACCTTACTCTGACAGCA 1920
QY 701 GlnSerGlnProHisGlyValTyrProProSerSerValAlaGlnGlyGlnSerGlnGly 720
Db 1921 CAGTCTACGCCCAAGGGGTATATCCACCTCAAGTGTGGCACAGGGGACAGACGAGGT 1980
QY 721 GlnProSerSerSerSerLeuThrGlyValSerSerSerGlnProIleGlnHisProGln 740
Db 1981 CAGCATCTCTCAAGTACGTTTAAACAGGGGTTTCACTTCCCAACCATACATCACTCTG 2040
QY 741 GlnGlnGlnGlyIleGlnGlnThrAlaProProGlnGlnThrValGlnTyrSerLeuSer 760
Db 2041 CAGCAGCAGGAAATACAGACAGACAGCCCTCCTCAACAGACAGTGCAGATTCATTCA 2100
QY 761 GlnThrSerThrSerSerGlnAlaThrThrAlaGlnProValSerGlnProGlnAlaPro 780
Db 2101 CAGACATCACTTCAGTGAAGGCCACTGACAGCCAGCAGAGTCACTCCACCTCA 2160
QY 781 GlnValLeuProGlnValSerAlaGlyLysGlnSerThrGlnGlyValSerGlnValAla 800
Db 2161 CAAGCTTGCCTCAAGTATCAGCTGGAAGAAACAGAGTACTCAGGAGTCTTCAAGTTCT 2220
QY 801 ProAlaGlnProValAlaValAlaGlnProGlnAlaThrGlnProThrThrLeuAlaSer 820
Db 2221 CCTGACAGGCGAGTGGACGTAGCAGACAGCCCAAGCTACCCAGCCGACACACTTGGCTTCC 2280
QY 821 SerValAspSerAlaHisSerAspValAlaSerGlyMetSerAspGlyLysGlnAsnVal 840
Db 2281 TCTGTAGACGTGACACATTCAGATGTGCTTCAGATGATGATGATGCAAGAGAAACGTC 2340
QY 841 ProSerSerSerGlyArgHisGlnGlyArgThrThrLysArgHisTyrArgLysSerVal 860
Db 2341 CCATCTTCCAGTGAAGGATGAGAGAAAGAACTCAAAACCGCATTCACGAAATCTGTA 2400
QY 861 ArgSerArgSerArgHisGlnLysThrSerArgProLysLeuArgIleLeuAsnValSer 880
Db 2401 AGGAGTGCCTCTCGACATGAAAAAACTTCAGCCCAAAATTAAGAAATTTGAATGTTCA 2460
QY 881 AsnLysGlyAspArgValValGlnCysGlnLeuGlnThrHisAsnArgLysMetValThr 900
Db 2461 AATTAAGAGACCGAGTAGTAGAATGTCAATTAAGACTCAATATAGGAAATGCTATCA 2520
QY 901 PheLysPheAspLeuAspGlyAspAsnProGlnGlnIleAlaThrIleMetValAsnAsn 920
Db 2521 TTCAAAATTTGACCTAGATGTGACAAACCCGAGAGAGTAGCAATATATGTGAAACAT 2580
QY 921 AspPheIleLeuAlaIleGlnArgGlnSerPheValAspGlnValArgGlnIleIleGln 940
Db 2581 GACTTATTTCTTACCAATAGAGAGAGAGTCTTGTGTGATCAAGTGCAGAAATTAATGAA 2640
QY 941 LysAlaAspGlnMetLeuSerGlnAspValSerValGlnProGlnGlnLysAspGlnGlnLeu 960
Db 2641 AAAGCTCATAAATGCTTCAAGTGAAGATGTCAAGTGAACCAAGGGGTGATCAAGGATG 2700
QY 961 GluSerLeuGlnGlnLysAspAspTyrGlyPheSerGlySerGlnLysLeuGlnGlnGln 980
Db 2701 GAGAGTCTACAGAGAAAGATGACTATGCTTTTCAAGGTTCTCGAATATGGAAGAGAG 2760
QY 981 PheLysGlnProIleProAlaSerSerMetProGlnGlnIleGlyIleProThrSerSer 1000
Db 2761 TTCAAACAAACAAATTCCTGTGCTTCCATGCCACAGCAAAATAGGACTTCACAGTTC 2820
QY 1001 LeuThrGlnValValHisSerAlaGlyArgArgPheIleValSerProValProGlnSer 1020
Db 2821 TTAACCTCAAGTTTATCATTTTGGCGGAGGGGTTTATATGAGTCTCTGCGCAAGAAAGC 2880
QY 1021 ArgLeuArgGlnSerLysValPheProSerGlnIleThrAspThrValAlaAlaSerThr 1040

Dh 2881 CGATTACGAGATCAAAAGTTTCCCGACGTGAATTAACATACATGTTGCTGCTTACA 2940
Qy 1041 AlagInserProGlyMeAsnLeuSerHisSerLeuSerLeuGlnGlnAla 1060
Db 2941 GCTAGAGCCCTGGAAATTAATTTGCTCCTGATCATCTTATGCTTCAACAGGCC 3000
Qy 1061 PheSerGluLeuAgaGalaGlnMeThrGluGlyProAsnThrAlaProProAsnPh 1080
Db 3001 TTTTCTGAACTTACGCGCCCAATATGACAGAGACCAATACAGCCTCCAAACTTT 3060
Qy 1081 SerHisThrGlyProThrPheProValValProPheLeuSerHisLeuGlyVal 1100
Db 3061 AGCATACAGAGACCAATTTCCAGTACCTCTTCTTAAGTACGATTCGAGGTC 3120
Qy 1101 ProThrThrAlaAlaAlaThrAlaProValProAlaThrsSerProProAsnAspLe 1120
Db 3121 CCAACACAGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 3180
Qy 1121 SerThrSerValIleGlnSerGluValThrValProThrGluGlnGlyIleAlaGlyVal 1140
Db 3181 TCCACATCAGTAATTCAGTCTGAGGTTACAGTCCCACTGAAGAGGGGATTCGAGATT 3240
Qy 1141 AlaThrSerThrGlyValValThrSerGlyGlyLeuProIleProProValSerGluSer 1160
Db 3241 GCCACGACGACAGGAGTGTGTACTTCAAGTGTCTCCCATACCACTGTGTCTGAATCA 3300
Qy 1161 ProValLeuSerSerValValSerSerIleThrIleProAlaValAlaSerIleSerThr 1180
Db 3301 CCAATACCTTCCAGCGTAGTTTCAAGTATCAATACCTGCAAGTTGTCTCAATATCTACT 3360
Qy 1181 ThrSerProSerLeuGlnValProThrSerThrSerGluIleValValSerSerThrAla 1200
Db 3361 ACATCCCGCTCACTTCAAGTCCCAATCATCATCATCATCATCATCATCATCATCATCAT 3420
Qy 1201 LeuThrProSerValThrValSerAlaThrSerAlaSerAlaGlyValSerThrAlaThr 1220
Db 3421 CTGATACCTTCAAGTCAAGTTCAGCACTTCAAGCTTCGAGGGGAGTACCTGATCC 3480
Qy 1221 ProGlyProLysProProAlaValValSerGlnGlnAlaAlaGlySerThrThrValGly 1240
Db 3481 CCAAGTCTTAAGCTCCAGCTGTAGTATCTCAGCAGGACGACGACGACGACGACGACG 3540
Qy 1241 AlaThrLeuThrSerSerValSerThrThrThrSerPheProSerThrAlaSerGlnLeuSer 1260
Db 3541 GCCACATTAACATGATGTTTCAACACCACTTCAATCCCAAGACAGCTTCACAGCTGCC 3600
Qy 1261 IleGlnLeuSerSerSerThrThrProThrLeuAlaGluThrValValValSerAla 1280
Db 3601 ATTCACTTACGACGACGATCTTCACTCTTACTTTCGTAACCGGTGTAGTTAGCCCA 3660
Qy 1281 HisSerLeuAspLysThrSerHisSerSerThrThrGlyLeuAlaPheSerLeuSerAla 1300
Db 3661 CACTACATGATTAACATCTCAAGACGATCAACCTGATGAGCTTCTCTCTCTCTCA 3720
Qy 1301 ProSerSerSerSerSerProGlyValAlaGlyValSerSerThrIleSerGlnProGlyIly 1320
Db 3721 CCATTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3780
Qy 1321 LeuHisProLeuValIleProSerValIleAlaSerThrProIleLeuProGlnAlaAla 1340
Db 3781 CTGCACTCTTGGTTCATTCATCACTGATGATGCTTCACTCTCTCTCTCTCTCTCTCT 3840
Qy 1341 GlyProThrSerThrProLeuLeuProGlnValProSerIleProProLeuValGlnPro 1360
Db 3841 GGAACCTACTTCAACCTTATTAACCCCAAGTACCTAGATCCCACTTGTGTACAGCT 3900
Qy 1361 ValAlaAsnValProAlaValGlnGlnThrLeuIleHisSerGlnProGlnProAlaLeu 1380
Db 3901 GTTGCCATATGCGCTGCTGACGACGACATTAATTCATATGACGCTCAACCACTTTG 3960
Qy 1381 LeuProAsnGlnProHisThrHisGlyProGluValAspSerAspThrGlnProLysAla 1400
Db 3961 CTTCCCAACCAAGCCCATATCTCATTTGCTGAAAGTATTCGATACCAACCAAGCT 4020

Qy 1401 ProGlyIleAspAspIleLysThrLeuGlnGlnLysLeuArgSerLeuPheSerGluHis 1420
Db 4021 CTTGCAATGATGACATTAAGACTCTGAAGAAAGCTGGCGTCTCTGTTCACTGAACAC 4080
Qy 1421 SerSerSerGlyValGlnHisAlaSerValSerLeuGlnThrSerLeuValIleGluSer 1440
Db 4081 AGCTCATCTGAGTCCGACATGCTCTGTCTCATCTGAGACCTCACTAGTCATGAGAGC 4140
Qy 1441 ThrValThrProGlyIleProThrThrAlaValAlaProSerLysLeuLeuThrSerThr 1460
Db 4201 ACAAGTACTTGCTTACACCAACCAATTTACATGAGACAGATTCCTTCCAGCTTACA 4260
Qy 1481 ProValValThrProGlyGlnValSerThrProValSerThrThrThrSerGlyValLys 1500
Db 4261 CCACTGTGTCAACCTGGGCAAGTTTCTAACCCAGTCAGCACTTACATCAGAGGTGAAA 4320
Qy 1501 ProGlyThrAlaProSerLysProProLeuThrLysAlaProValLeuProValGlyThr 1520
Db 4321 CTTGAACTGCTCTCTCTCAAGCCACTTCACTTAAGCTCCGCTGCTGCCAGTGGTACT 4380
Qy 1521 GluLeuProAlaGlyThrLeuProSerGlnGlnLeuProPheProGlyProSerLeu 1540
Db 4381 GAATCTTCAGAGGTACTCTACCCAGCAGAGCTGCCACTTTCCAGACCTTCTCTA 4440
Qy 1541 ThrGlnSerGlnGlnProLeuGlnAspLeuAspAlaGlnLeuAgaGlyThrLeuSerPro 1560
Db 4441 ACCAGTCCCGACCACTTCAAGATCTGTATCTCAATTAAGAAAGAACTTAAGTCA 4500
Qy 1561 Glu**IleThrValIleThrSerAlaValAlaGlyProValSerMetAlaAlaProThrAlaIle 1580
Db 4501 GAGATATCAGAGGACTTCTGCGGTGTGCTGTGCTCATGGGGCTCCAAACAGCATTC 4560
Qy 1581 ThrGlnAlaGlyThrGlnProGlnLysGlyValSerGlnValLysGlnGlyProValLeu 1600
Db 4561 ACAGAGAGAGAACACAGCTCAGAGAGGTCTTCTCAAGTCAAAAGAGGCCCTGTCTTA 4620
Qy 1601 AlaThrSerSerGlyValAlaGlyValPheLysMetGlyArgPheGlnValSerValAlaAla 1620
Db 4621 GCAACTAGTTCAGAGGCTGTGTTTAAAGTGGACCATTTCAAGTTTCGTTGTCACCA 4680
Qy 1621 AspGlyAlaGlnLysGlnGlyLysAsnLysSerGlnAspAlaLysSerValHisPheGlu 1640
Db 4681 GACGATGCCGAAAGAGGGTAAATTAAGTCAAGATGCAAAAGTCTGTTCAATTGAA 4740
Qy 1641 SerSerThrSerGlnSerSerValLeuSerSerSerSerProGlnSerThrLeuValLys 1660
Db 4741 TCCAGACCTCAGAGTCTCTCACTCATCAAGTATGATCCAGAGATCACTTGGTGA 4800
Qy 1661 ProGluProAsnGlyIleThrIleProGlyIleSerSerAspValProGlnSerAlaHis 1680
Db 4801 CCAGAGCCGAATGCAATACCATCCCTGATTCCTTCAAGTGTGCCAAGAGTCCAC 4860
Qy 1681 LysThrThrAlaSerGlnAlaLysSerAspThrGlnGlnProThrLysValGlyArgPhe 1700
Db 4861 AAATCTACTGCTTCAAGGCAAGTCAAGACTGGGACGCTTCAACAGTTGGAGCTTT 4920
Qy 1701 GlnValThrThrThrAlaAsnLysValGlyArgPheSerValSerLysThrGluAspLys 1720
Db 4921 CAGGTGACATCAGCAAAACAAAGTGGTGTCTCTGATCAAAACAGAGAGCAAG 4980
Qy 1721 IleThrAspThrLysGlyGlnGlyProValAlaSerProPhePheAspLeuGlnGln 1740
Db 4981 ATCACTGACAAAGAAAGAGAGACAGTGCATCTCTCTTATGATTTGAAACAA 5040
Qy 1741 AlaValLeuProAlaValIleProLysLysGlnLysProGlnLeuSerGluProSerHis 1760
Db 5041 GCTGTTCTTCTGCTGTATATCAAAAGAAAGAGAGCTGAAGCTCAAGAGCTTCAAT 5100

QY	1761	LeuAnngLYpProSeSerSeShpProGlu1aaiApeLeuSeSerAArgAspValAspAspGly	1780
Db	5101	CTAAATGGGCGCCTCTTCGACCCTGGAGGCGCGCTTTTAAAGTAGGGAAGTGAATGATGCT	5168
QY	1781	SerGlySerProHisSerProHisGlnLeuSerSerLySerLeuProSerGlnAsnLeu	1800
Db	5161	TCCGGTAGTCCACACTCGCCCATAGCTGAGCTGAAGAGCGCTTCTAGCCAGAAATCTA	5220
QY	1801	SerGlnSerLeuSerAsnSerPheLeuSerSerTyMetSerSerAspAsnGluSerAsp	1820
Db	5221	AGTCAAAAGCTTAGTAATTCATTTCAATCTCTCTTCATGAGTAGAGCAATATGAGTCAGAT	5280
QY	1821	ILegIuaSpGLuaSpLeuLyLeuGluLeuLeuArgAArgLeuAArgAspLyHisLeuLyGlu	1840
Db	5281	ATCGAAGATGAAGACTTAAGTTAAGCTGCACACACACTACAGATTAACATCTCAAAAG	5340
QY	1841	ILegIuaSpLeuGlnSerArgGlnLyHisGluILegIuaSerLeuTyThLysLeuGly	1860
Db	5341	ATTGAGAGCTGCACAGAGTCGCAGAGAGATGAATTTGAATCTTTGTATACAAACTGGGC	5400
QY	1861	LysValProProAlaValILeILeProProAlaAlaProLeuSerGlyArgArgArgArg	1880
Db	5401	AAGGAGCCCCCTGCTGTTATTTATTCGCCACCTGCTCCCTTCAGGAGAGAAGCAGCA	5460
QY	1881	ProThLySerSerGlySerLySerSerArgSerSerSerSerLeuGlyAsnLySerPro	1900
Db	5461	CCCACTTAAGAAGCAAGCAAGCAAACTGATGGAAGCACTTCTTGGGAAATTAAGCCCC	5520
QY	1901	GlnLeuSerGlyAsnLeuSerGlyGlnSerAlaAlaSerValLeuHisProGlnGlnThr	1920
Db	5521	CAGCTTTCAAGTAACCTGTCTGTGCAGAGTGCAGCTTCAGCTTGACCCCCCAGCAAGCC	5580
QY	1921	LeuHisProProGlyAsnILeProGluSerGlyGlnLeuGlnLeuGlnProLeuLys	1940
Db	5581	CTCCACCCCTCGGCAACATCCCAAGAGTCCGGGCAAGATCAGCTGTATACAGCCCTTAAG	5640
QY	1941	ProSerProSerSerAspAsnLeuTySerAlaPheThrSerAspGlyAlaILeSerVal	1960
Db	5641	CAATCTCCCTCAGAGACACCTCTATTCACCTTCACCACTGATGAGGCCATTTCCAGTA	5700
QY	1961	ProSerLeuSerAlaProGlyGlnGlyThrSerSerThraSerThrValGlyAlaThrVal	1980
Db	5701	CCAAAGCCTTTCTGCTCCAGGTCGACAGAAACGACGACAAACACTGTTGGGCAACAGTG	5760
QY	1981	AsnSerGlnAlaAlaGlnAlaGlnProProAlaMetThrSerSerArgLySgLYThrPhe	2000
Db	5761	AACAGCCAAAGCGCCCAAGCTCAGCTCTCTCCATGACGTCCACAGAGAAAGGCACTTC	5820
QY	2001	ThraPheAspLeuHisLySgLYLeuValAspAsnTPAlaArgAspAlaMetAsnLeuSerGly	2020
Db	5821	ACAAGATGACTTGACAAAGTTGTGTGACAAATTTGGGCCGAGAGTGCATGAAATCTCTCAGGC	5880
QY	2021	ArgArgGlySerLySgLYHisMetAsnTyGlnGlyProGlyMetAlaArgLySgLYPheSer	2040
Db	5881	AGGAGAGGAACCAAGGCACTGAATTAATGAGAGGCCCTTGAAATGGGAAAGGAATTTCT	5940
QY	2041	AlaProGlyGlnLeuCyHisLeuSerMetThraSerAsnLeuGlyGlySerAlaProILeSer	2060
Db	5941	GCACCTGGGGCAACTGTGCATCTCCATGACCTCGAACTCGGAGTGGCTGTGCCCCCACTCT	6000
QY	2061	AlaAlaSerAlaThrThrSerLeuGlyHisPheThrLySerMetCySProProGlnGlnTyR	2080
Db	6001	GCACACATCAGTACTCTCTTAGGCACTTCAACAAAGCTCTATGTGCCCCCCCAAGAGATAT	6060
QY	2081	GlyPheProAlaThrProPheGlyAlaGlnTySerGlyThrGlyGlyProAlaProGln	2100
Db	6061	GGCTTTCAGCTACCCCAATTTGGGCTCAATGAGAGTGGAGAGGGTGGCCCAAGCAACAG	6120
QY	2101	ProLeuGlyGlnPheGlnProValGlyThraLaserLeuGlnAsnPheAsnILeSerAsn	2120
Db	6121	CCAACTGGCCAGTTCCAAACCTGTGGGAAGTGGCTCTTGAGATTTCAACATCACAGCAAT	6180
QY	2121	LeuGlnLySerILeSerAsnProProGlySerAsnLeuArgThrThr	2136

[illegible]

QY 141 AlaAlaProGluGluGlnAlaValAlaGlyProAlaProSerThrValProSerSerThr 160
DB 241 GCCGCCCTGGGGAAACAGCCGTCGCGGGCCCTGCCCTCGACTGTCCTCCAGAGTACC 300
QY 161 SerIysAspArgProValSerGlnProSerLeuValGlySerIysGluGluProProPro 180
DB 301 AGCAAAACCCGCCAGTGTCCAGCCCTGCTGTGGGAGCAAAAGAGAGCCGCCGCCG 360
QY 181 AlaArgSerGlySerGlyGlySerAlaIysGluProGlnGluIysArgSerGln 200
DB 361 GCGAAGATGGCAGCGCGCGCGAGCCCAAGAGCCACAGAGAAACGAGCCAGCAG 420
QY 201 GluAspAspIleGluGluLeuGluThrLysAlaValGlyMetSerAsnAspGlyArgPhe 220
DB 421 CAGAGTATATCGAAAGCTGGAGACCAAGGCCCGGGGAATGTCTAAAGATGCCCTTT 480
QY 221 LeuLysPheAspIleGluIleGlyArgGlySerPheLysThrValTyrLysGlyLeuAsp 240
DB 481 CTCAAGTTTGCATCGAAATCGCAGAGGCTCTTTAAGACGGTCTCAAAAGGCTGAGC 540
QY 241 ThrGluThrThrValGluValAlaTyrProGluLeuGluAspArgLysLeuThrLysSer 260
DB 541 ACTGAACCCAGCCGAGAGTCCTCGTGTGTAATCGACGATCCAAATTTAACAAAGTCT 600
QY 261 GluArgGluArgPheLysGluGluAlaGluMetLeuLysGlyLeuGlnHisProAsnIle 280
DB 601 GAGAGCGAGAGATTTAAAGAAAGAGCTGAATGTAAAGAGTCTTCAGCATCCCAATTT 660
QY 281 ValArgPheLysArgPheSerTyrGluSerThrValLysGlyLysLysGlyLeuVal 300
DB 661 GTTAGATTTTATGATTTCTCTGGAAATCCACAGTAAAGAAAGAAAGAGTATGTTGGTG 720
QY 301 ThrGluLeuMetThrSerGlyThrLeuLysThrTyrLeuLysArgPheLysValMetLys 320
DB 721 ACTGAATTAATGAGCTGTGAACACTTAAACGTAATCGAAAGAGTTTAAAGTATGAG 780
QY 321 IleLysValIleArgSerTyrCysArgGlnIleLeuLysGlyLeuGlnPheLeuHisThr 340
DB 781 ATCAAGTCTTAAGAGAGTGGTGGCGGTGAGATCCCTTAAGGCTTCCTTCATCTACT 840
QY 341 ArgThrProProIleIleHisArgAspLeuLysCysAspAsnIlePheIleThrGlyPro 360
DB 841 CGAATCCACCTATCATTCACCGGATCTTAATGTGCAACATCTTATCAACCGGCCCT 900
QY 361 ThrGlySerValLysIleGlyAspLeuGlyLeuAlaThrLeuLysArgAlaSerPheAla 380
DB 901 ACTGGCTAGTCAGATTTGAGAGACTCGGTCTGGCAACCTCGAGCGGGCTTCCTTTCC 960
QY 381 LysSerValIleGlyThrProGluPheMetAlaProGluMetTyrGluGluLysTyrAsp 400
DB 961 AAGAGTGTATAGTATGCCAGAGTTCATGGCCCTGAGATGATGAGAGAAATATGAT 1020
QY 401 GluSerValAspValTyrAlaPheGlyMetCysMetLeuGluMetAlaThrSerGlyTyr 420
DB 1021 GAATCCGTTGACGTTATGCTTTTGGAGTGTGATGCTTGAAGTGGCTACATCTGAATAT 1080
QY 421 ProTyrSerGlyCysGlnAsnAlaIleGlnIleTyrArgArgValThrSerGlyValLys 440
DB 1081 CTTTACTGGAGTGCAGAAATGCTGACAGATTCACCGTGGCTGACCAAGTGGGGTAAAG 1140
QY 441 ProAlaSerPheAspLysValAlaIleProGluValLysGluIleIleGluGlyCysIle 460
DB 1141 CCAGCCAGTTTGAACAAAGTATTCCTGAAAGTGAAGAAATATATGAAAGATGACATA 1200
QY 461 ArgGluAsnLysAspGluArgTyrSerIleLysAspLeuLeuAsnHisAlaPhePheGln 480
DB 1201 CGACAAACCAAGATGAAAGATATTCATCAAAAGACCTTTGAACCATGCCCTTCTCCA 1260
QY 481 GlnGluThrGlnLysValArgValGluLeuAlaGluGluAspAspGlyGluLysIleAlaIle 500
DB 1261 GAGGAAACGAGGTACGGGTATGAGCAAGAAAGATGATGAGAAATTAATGACCATTA 1320
QY 501 LysLeuTyrLeuArgIleGluAspIleLysLysLysLysLysLysLysLysLysLys 520

DB 1321 AAATTTAGCTACGATATTGAAGATATTAAAGAAATTAAGGAAATTAAGAAATATGAA 1380
QY 521 AlaIleGluPheSerPheAspLeuGluArgAspValProGluAspValAlaGlnGluMet 540
DB 1381 GCTATTGACTTTTCTTTTGTATTAGAGAGAGATGTCCCAAAAGATGTTCACAAAGAAATG 1440
QY 541 ValGluSerGlyTyrValCysGluGlyAspHisLysThrMetAlaLysAlaIleLysAsp 560
DB 1441 GTAGAGTCTGGATATGCTGTGAAGGTGATCAACAGCATGCTTAAGCTATCAAAAGAC 1500
QY 561 ArgValSerLeuIleLysAlaGlyArgGluGluArgGluMetValArgGluGluGlu 580
DB 1501 AGAGTATCATTAATTAAAGAGAAACGAGACAGCGGAGTGTGTGACGGAGAGCAAGAA 1560
QY 581 LysLysLysGlnGluGluSerSerLeuLysGlnValGluGlnSerAlaSerGln 600
DB 1561 AAAAAAAGCAGAGAGAGAGAGAGCTCAAAACAGAGAGTAAACATCAATCAATGCTTCCAG 1620
QY 601 ThrGlyIleLysGluLeuProSerAlaSerThrGlyIleProThrAlaSerThrThrSer 620
DB 1621 AAGAGATCAAGCAGCTCCCTTCTGTAGCAGCGCATACTACTGCTTCAACCATTTCA 1680
QY 621 AlaSerValSerThrGlnValGluProGluGluProGluAlaAspGlnHisGlnGluLeu 640
DB 1681 GCTTCAGTTCTTACCAAGTAGAACTGAAGAACTGAGGAGATCAACATCAACAACTA 1740
QY 641 GlnTyrGlnGlnProSerIleSerValLeuSerAspGlyThrValAspSerGlyGlnGly 660
DB 1741 CAGTACCAAGCAACCAATATATCTGTATTCTATGAGGAGTTCAGAGTGTCAAGGGA 1800
QY 661 SerSerValPheThrGluSerArgValSerSerGlnGlnThrValSerTyrGlySerGln 680
DB 1801 TCCTCTGCTTCAAGATCTCGATGAGGAGCCAAAGACAGTTTCATATGTTCCCAA 1860
QY 681 HisGluGlnAlaHisSerThrGlyThrValProGlyHisIleProSerThrValGlnAla 700
DB 1861 CATGAACAGGACATCTTACAGCAGCAGCTCCAGGAGATATACCTTCACTGCTCAACCA 1920
QY 701 GlnSerGlnProHisGlyValTyrProProSerSerValAlaGlnGlyGlnSerGlnGly 720
DB 1921 CAGTCTAGCCCAATGGGATATTCACCTCAAGTGTGACAGGAGGACAGCCAGGCT 1980
QY 721 GlnProSerSerSerSerLeuThrGlyValSerSerSerGlnProIleGlnHisProGln 740
DB 1981 CAGCATCTCAAGTATGATTAAGGGGTTTCATCTCCCAACCAATACATCTCTGAG 2040
QY 741 GlnGlnGlnGlyIleGlnGlnThrAlaProProGlnGlnThrValGlnTyrSerLeuSer 760
DB 2041 CAGCAGCAGGAGAAATACAGCAGACAGCCCTCTCAACAGACAGTGCAGTATTCACCTTCA 2100
QY 761 GlnThrSerThrSerSerGluAlaThrThrAlaGlnProValSerGlnProGlnAlaPro 780
DB 2101 CAGACATCACTCAAGTACGAGGCGCATCTACACAGCCAGTGAAGTCAAGCTCAAGCTCA 2160
QY 781 GlnValLeuProGlnValSerAlaGlyLysGln 791
DB 2161 CAAGTCTTGCTCAAGTATCAGTGTGAAAAACAGGGCTTCCCACTCGATCGACACAGAG 2220
QY 791 ----- 791
DB 2221 TACCCAGAGATTCAAATATTTGCTCCCTTCCAAAGTGGCTTCTGTTGATCCATCT 2280
QY 791 ----- 791
DB 2281 ACAGTCTAAKCCCTCCATCGCCAGACAGAAATACGTGCTACACCTGGGTATCTTCCCA 2340
QY 791 ----- 791
DB 2341 GTGTGAGCCTTATGTGAATCAAACTTTTATGTTCTATGGGTGTGTAGAGAGAGAG 2400
QY 791 ----- 791

Dh 2401 GTTCAGTGTCCACGACGAGGAGGATTGACAAAGCCCACTACCTCTCCAGCAA 2460
Qy 792 -----SerThrGlnGlyValSerGlnValAlaProAlaGlnProValAlaVal 807
Db 2461 GCAGTTTGGAGAGTACTCAGGAGAGTCTTCAGGTTCCTCCGACGAGCAATGGAGTA 2520
Qy 808 AlaGlnProGlnAlaThrGlnProThrThrLeuAlaSerSerValAspSerAlaHisSer 827
Db 2521 GCACAGCCCCAGACTACCCAGCAGCACTTGGCTCTCTCTGTMGACAGTCCACTTCA 2580
Qy 828 AspValAlaSerGlyMetSerSerGlyValGlnValSerValProSerSerSerGlyValHis 847
Db 2581 GATGTGCTTCAGGATAGATGATGCAATGAAAGCTCCACTCTCCAGTGAAGGCAAT 2640
Qy 848 GluGlyArgThrThrLysArgHisTyrArgLysSerValArgSerArgHisGln 867
Db 2641 GAAGGAAAGAACTACAAAGCGATTACGAAATCTGTAAGAGTCGCTTCGACATGAA 2700
Qy 868 LysThrSerArgProLysLeuArgLysLeuValSerAsnLysGlyAspArgValVal 887
Db 2701 AAAACTCAGCCCAAAATTAAAGATTGTAATGTTCAATTAAGAGAGACCGAGTAGTA 2760
Qy 888 GluGlyGlnLeuGlnThrHisAsnArgLysMetValThrPheLysPheAspLysGly 907
Db 2761 GAATGTCAATTAGAGACTCATATATAGAAATGTTACATTCMAATTTGACCTAGATGCT 2820
Qy 908 AspAsnProGlnGlnLeuAlaThrIleMetValAsnAsnAspPheIleLeuAlaIleGln 927
Db 2821 GACAACCCCGAGAGATAGCAACATTTATGGAAACAAATGACTTTATCTTGCAATTAAG 2880
Qy 928 ArgGlnSerPheValAspGlnValArgGlnIleIleGlnLysAlaAspGlnMetLeuSer 947
Db 2881 AGAGAGTGTGTTGTGATCAAGTGCAGAAATATTATTAAGAAAGCTGATGAATGTGCT 2940
Qy 948 GluAspValSerValGlnProGlnGlyAspGlnGlyLeuGlnSerLeuGlnGlyLysAsp 967
Db 2941 GAGAGTGCAGTGTGAGAACCAAGAGGTGATCAGGATTTGAGAGTCTCAAGAGAAAGAT 3000
Qy 968 AspTyrGlyPheSerGlySerGlnLysLeuGlnGlyGlnPheLysGlnProIleProAla 987
Db 3001 GACTATGGCTTTTCAGGTCTTCAGAAATTTGAAAGAGAGTTCMAACAACTATTCCTGCG 3060
Qy 988 SerSerMetProGlnGlnIleGlyLysLeuProThrSerSerLeuThrGlnValHisSer 1007
Db 3061 TCTTCACATGCCACAGCAATAGGCAATTCCTACAGTTCTTTACTCAAGTGTTCATCT 3120
Qy 1008 AlaGlyArgArgPheIleValSerProValProGlnSerArgLeuAlaGlnSerLysVal 1027
Db 3121 GCGGGAAGGCGGTTTATAGTAGTCTGTGTCAGAAAGCCGATTAACGAAATCAAAAGTT 3180
Qy 1028 PheProSerGlnIleThrAspThrValAlaAlaSerThrAlaGlnSerProGlyMetAsn 1047
Db 3181 TTCCCGATGAAATACAGATACAGTACGTTCTCTTACAGTCCAGAGCCCTGGAGTAGAC 3240
Qy 1048 LeuSerHisSerAlaSerSerLeuSerLeuGlnGlnAlaPheSerGlnLeuArgArgAla 1067
Db 3241 TTGTCTCACTCTGCATCATCCCTAGTCTACCAACAGGCTTTTCGAACCTTAAGCGTCC 3300
Qy 1068 GluMetThrGlnGlyProAsnThrAlaProProAsnPheSerHisThrGlyProThrPhe 1087
Db 3301 CMAATGACAGAGAACCCAAACAGACCTCCAAACTTTAGTCTTACAGGACCAACATTT 3360
Qy 1088 ProValValProProPheLeuSerSerIleAlaGlyValProThrThrAlaAlaThr 1107
Db 3361 CCAGTACTACTCTTCTTTAAGTAGCATGCTGAGATCCCAACACAGACGACGCCACA 3420
Qy 1108 AlaProValProAlaThrSerSerProProAsnAspIleSerThrSerValIleGlnSer 1127
Db 3421 GCACCACTCCCTGCACAAAGAGCCCTCTAATGATGATTCACATCATGATATCACTCT 3480
Qy 1128 GluValThrValProThrGlnGlnGlyIleAlaGlyValAlaThrSerThrGlyValVal 1147
Db 3481 GAGGTTACAGTGTCCCACTGAAAGAGGAGATTGCTGAGATTGCCACAGACAGAGTGTGTA 3540

Qy 1148 ThrSerGlyGlyLeuProIleProProValSerGlnSerProValLeuSerSerValVal 1167
Db 3541 ACTTCAGGTGTCTCCATACCACTGTGTGATCATCACAGTACTTCCAGCTAGATT 3600
Qy 1168 SerSerIleThrIleProAlaValValSerIleSerThrThrSerProSerLeuGlnVal 1187
Db 3601 TCMAATATCACAAATCTCGCACTTCTTCATATATCTACATATCCCGCTCACTCAAGTC 3660
Qy 1188 ProThrSerThrSerGlnIleValValSerSerThrAlaLeuThrProSerValThrVal 1207
Db 3661 CCCACATCCACATCTGAGATGTTGTTCTAGTACAGACATGTATCTTCACTACATACAGTT 3720
Qy 1208 SerAlaThrSerAlaSerAlaGlyGlySerThrAlaThrProGlyProLysProProAla 1227
Db 3721 TCAGCACTTCACTCTGACAGGGGGAGTCTGTACCCAGAGTCTTAAGCTTCACACT 3780
Qy 1228 ValValSerGlnGlnAlaAlaGlySerThrThrValAlaThrLeuThrSerValSer 1247
Db 3781 GTAGATCTCAGCAGCAGCAGCAGCAGCTACTGTGGAGCCACTTAACATCAAGTTCT 3840
Qy 1248 ThrThrThrSerPheProSerThrAlaSerGlnLeuSerIleGlnLeuSerSerThrThr 1267
Db 3841 ACCACACTTCATTCACCAAGCAGCTTCACAGCTGTCCATTCAGCTTAGCAGAGTACT 3900
Qy 1268 SerThrProThrLeuAlaGlnThrValValValSerAlaHisSerLeuAspLysThrSer 1287
Db 3901 TCTACTCTTACTTATGAGAAACCGGTGATGATGCGCACACTCATGATTAAGACATCT 3960
Qy 1288 HisSerSerThrThrGlyLeuAlaPheSerLeuSerAlaProSerSerSerSerPro 1307
Db 3961 CATAGCACTAACAATGATTTGGCTTTCTCTCTCTGCAACATCTTCTCTCTCTCTCT 4020
Qy 1308 GlyAlaGlyValSerSerTyrIleSerGlnProGlyGlyLeuHisProLeuValIlePro 1327
Db 4021 GGAGCAGAGAGTGTAGTTATATTTCTCAGCCTGTGGGTGTCATCTTGTGATTTCA 4080
Qy 1328 SerValIleAlaSerThrProIleLeuProGlnAlaAlaGlyProThrSerThrProLeu 1347
Db 4081 TCAATGATAGCTTCTACTCTCATTTCTTCCCAACAGAGAGACCTTCTTCAACCTTTA 4140
Qy 1348 LeuProGlnValProSerIleProProLeuValGlnProValAlaAsnValProAlaVal 1367
Db 4141 TTACCCCAAGTACTATGATATCCACCTTGGTACAGCTGTGCAATGTGCTGCTGTA 4200
Qy 1368 GlnGlnThrLeuIleHisSerGlnProGlnProAlaLeuLeuProAsnGlnProHisThr 1387
Db 4201 CAGCAGACACTAATTCATATGACAGCTCAACCACTTGTGCTTCCCAACCAAGCCCATACT 4260
Qy 1388 HisGlyProGlnValAspSerAspThrGlnProLysAlaProGlyIleAspAspIleLys 1407
Db 4261 CATGTCTTGAAGTAGATTCTGATACAAACCCAAAGCTCTGTGAATTTGATGACATTAAG 4320
Qy 1408 ThrLeuGlnGlnLysLeuArgSerLeuPheSerGlnHisSerSerSerGlyValAlaHis 1427
Db 4321 ACTCTAAGAAAAAGCTGCGGTCTCTGTTCAGTAAACACACTATCTGAGCTCAGAT 4380
Qy 1428 AlaSerValSerLeuGlnThrSerLeuValIleGlnSerThrValThrProGlyIlePro 1447
Db 4381 GCCCTGTCTCAGTGCAGAGACTCATGATGAGAGCACTGTACACAGCAGGATCCCA 4440
Qy 1448 ThrThrThrAlaAlaProSerLysLeuLeuThrSerThrThrSerThrCysLeuProPro 1467
Db 4441 ACTAGTGTGTGACCAAGCAAACTCTGATCTTCAACCAACTACTCTGTTACACACA 4500
Qy 1468 ThrAsnLeuProLeuGlnThrValAlaLeuProValThrProValValThrProGlyGln 1487
Db 4501 ACCAATTTACACATAGAAACAGTTGCTTTCAGTTACACAGTGTGACACACTGGAGCA 4560
Qy 1488 ValSerThrProValSerThrThrThrSerGlyValLysProGlyThrAlaProSerLys 1507
Db 4561 GTTCTTACCCCACTGACACTACTACTACATCAGAGTGAACCTGGAAGTGTCTCCCAAG 4620

Db 1621 ACAGAAATCAAGACGCTCCCTCTGTGCTAGACACCGGCACTACCTGCTTCTACCACTTCA 1680
QY 621 AAservalSerThrglnValGluProGlnGluProGlnAlaAspGlnHisGlnGlnLeu 640
Db 1681 GCTTCAGTTTCTTACACAGTAAGAACTGAAAGAACTTAGGCAATCAACATCAACAATCA 1740
QY 641 GlnTyrglnGlnProSerHisSerValIleuSerAspGlyThrValaAspSerGlyGlnGly 660
Db 1741 CAGTACACGACACCCAGTATATCTGTGTATCTGATGGACGGTTGACAGTGGTCAAGGGA 1800
QY 661 SerSerValPheThrGlnSerArgValSerSerGlnGlnThrValSerTyrglySerGln 680
Db 1801 TCCTCTGCTTCAACAGATCTCGAGTACGACGACCAAGACAGATTCATATGTGTCCAA 1860
QY 681 HisGlnGlnAlaHisSerThrglyThrValProGlnHisIleProSerThrValGlnAla 700
Db 1861 CATGAACAGGACATTTCTACAGGACAGTCCAGGGCATATACCTTCTACTGTCCAGCA 1920
QY 701 GlnSerGlnProHisGlyValTyProProSerSerValAlaGlnGlyGlnSerGlnGly 720
Db 1921 CAGTCTAGCCCACTGGGGTATATCCACCCTCAAGTGGGCAAGGGGCAAGGCGT 1980
QY 721 GlnProSerSerSerSerLeuThrGlyValSerSerSerGlnProIleGlnHisProGln 740
Db 1981 CAGCCATCTCTCAAGTACCTTAAACAGGGGTTTCATCTTCCCAACCAACAACATCTTCAG 2040
QY 741 GlnGlnGlnGlyIleGlnGlnThrAlaProProGlnGlnThrValGlnTySerLeuSer 760
Db 2041 CAGAGGACGGGAATACAGCACAGACCCCTCTTCAACAGACAGTGCAGTATTCCTTCA 2100
QY 761 GlnThrSerThrSerSerGlnAlaThrThrAlaGlnProValSerGlnProGlnAlaPro 780
Db 2101 CAGACATCAACCTCCAGTGAAGGCCATATCTGACACAGCAGAGTACAGCTTCAAGTCCA 2160
QY 781 GlnValIleuProGlnValSerAlaGlyLysGln----- 791
Db 2161 CAAGCTTGTGCTCAAGTATCAGCTGGAACAAACAGCTTCCAGTTTCCAGCCAGTACCACT 2220
QY 791 ----- 791
Db 2221 ATCCAGGCGAACCCTCAGATCCAGTTCGACACAAACCTCGGTTGTTCACGTCACCTCT 2280
QY 791 ----- 791
Db 2281 GGTGTCTATTTCTTCCAGTGGAGAGCGGCTCCCTACTCCCTGCTCCCTCAGTACCTCT 2340
QY 791 ----- 791
Db 2341 GTCTCTCAGATTCCCATATCAATCTTCATGTGTCTACGGCTCAGACAGAGTTTCTCATTC 2400
QY 791 ----- 791
Db 2401 CTTCCTCATCAAAATGAGCTGGCACTTACTAGCCTCTGCTCAGCTTGGCTTCATCTGCT 2460
QY 791 ----- 791
Db 2461 ACAACAGCTGGGATCCCGGGGTATCAACTGTGTCTTCTAGTCAAGCTTCAACCTTCTG 2520
QY 791 ----- 791
Db 2521 CAGCTGTGACTCAGCTGCAAGTCAAGTTCACCCACAGCTCTTCAACACAGAGTTGAG 2580
QY 791 ----- 791
Db 2581 TTCATGGGAATACAGCTTAACCTTGGACAAGCTGCTGAGGTTCACTTCTCTGAGAGAT 2640
QY 791 ----- 791
Db 2641 GTTCTGTACAGGGGCTTCCCACTGACAGTCCACACAGTACCCAGAGATTCATAATAT 2700
QY 791 ----- 791
Db 2701 GCTCCCTCTTCAACGTTGCTTGTGATCATCTTACAGTCTTACCCCTCCCATG 2760

QY 791 ----- 791
Db 2761 CCGACAGAAAGTACTGGCTACACCTGGGTACTTTCCACAGTGGTGACCTTATGTGAA 2820
QY 791 ----- 791
Db 2821 TCAATCTTGTAGTCTTATAGGTGTGTAGAGACAGGTCAAGTGTCCACGACGA 2880
QY 792 -----SerThrGln 794
Db 2881 GGGAGTTTACACAGCCCCCACTTACATCTCCACAGACAGCTTTGGAGAGTACTCAG 2940
QY 795 GlyValSerGlnValAlaProAlaGluProValAlaValAlaGlnProGlnAlaThrGln 814
Db 2941 GAGTCTCTCAGGTGTCTCTGACAGAGCAGTTGAGTATACACAGCCCCCAAGTACCCAG 3000
QY 815 ProThrThrLeuAlaSerSerValaAspSerAlaHisSerAspValAlaSerGlyMetSer 834
Db 3001 CCGACCACTTGGCTTCTCTGTGTAGACAGTGCACATTCAGATGTGTCTTCAAGTATAGT 3060
QY 835 AspGlyAsnGlnAsnValProSerSerSerGlyArgHisGlnGlyArgThrThrLysArg 854
Db 3061 GATGGCATAGAAACGTCCTCATCTTCCAGTGAAGGCAATGAAGAACTACAAAAACG 3120
QY 855 HisTyraArgLysSerValaArgSerArgSerArgHisGlnTyThrSerArgProLysIleu 874
Db 3121 CATACCGAAATCTGTAAAGAGTCGCTTCGACATGAAAAAATTCAAGCCCAAAATTA 3180
QY 875 ArgIleuLeuAsnValSerAsnLysGlyAspArgValAlaGlyCysGlnIleuGlnThrHis 894
Db 3181 AGATTTTGAATGTTTCAATTAAGAGACCGAGTATAGATCAATTAAGAACTCAT 3240
QY 895 AsnArgLysMetValThrPheLysPheAspLeuAspGlyAspAsnProGlnGlnIleAla 914
Db 3241 AATGGAAATGTGTTAATTCATTAATTTGACTTGAATGTGACAAACCCGAGGAATAGCA 3300
QY 915 ThrIleMetValAsnAsnAspPheIleLeuAlaIleGlnArgGlnSerPheValaAspGln 934
Db 3301 ACATTTATGTGAAACATGACTTATTCTGACATAGAGAGAGTGGTTGTGATCA 3360
QY 935 ValArgGlnIleIleGlnLysValaAspGlnMetLeuSerGlnAspValSerValGluPro 954
Db 3361 GTGGAGAAATTAATTAAGAAAAAGCTGATGAATATCTCAGTGGATGTCTAGTGGAAACA 3420
QY 955 GlnGlyAspGlnGlyLeuGlnSerLeuGlnGlyLysAspAspTyrglyPheSerGlySer 974
Db 3421 GAGGGTATACGGGATTTGAGAGCTTCAAGGAAAGATGACTATGGCTTTTCAGGTTCT 3480
QY 975 GlnLysLeuGlnGlyGluPheLysGlnProIleProAlaSerSerMetProGlnGlnIle 994
Db 3481 CAGAAATTTGAAAGAGAGTTCAAAACAACCAATTCCTGGGTCTTCATGCCACACAATA 3540
QY 995 GlyIleProThrSerSerLeuThrGlnValHisSerAlaGlyArgArgPheIleVal 1014
Db 3541 GGCAATCTCTACAGTTCTTTAACTCAAGTGTATCTCTGGGAGAGCGGTTTATATGTG 3600
QY 1015 SerProValProGlnSerArgLeuArgGlnSerLysValPheProSerGlnIleThrAsp 1034
Db 3601 AGTCTGTGCAAGAGCCGATTCAGAGATTCAAAAGTTTCCCAAGGAAATTAACAGAT 3660
QY 1035 ThrValAlaAlaSerThrAlaGlnSerProGlyMetAsnLeuSerHisSerAlaSerSer 1054
Db 3661 ACAGTGTGCTCTACAGCTCAGAGCCCTGGAATGAATCTGTCTACCTGTGACATATCC 3720
QY 1055 LeuSerLeuGlnGlnAlaPheSerGlyLeuArgArgAlaGlnMetThrGlnGlyProAsn 1074
Db 3721 CTTAGTCTACAAACGGGCTTTTCTGAACCTTGAAGCTGCCAAATGACGAAGAACCAAY 3780
QY 1075 ThrAlaProProAsnPheSerHisThrGlyProThrPheProValAlaProProPheLeu 1094
Db 3781 ACAGCACTCCAAACTTATAGTATACAGGACCAACATTTCCAGTATGATCCTCTTCTTA 3840

QY 1095 SerSerIleAlaGlyValProThrThrAlaAlaAlaThrAlaProValProAlaThrSer 1114
DB 3841 AGTAGCATTTGTGAGTCCCAACACAGACAGACACAGCTCCCTGCACAAAGC 3900
QY 1115 SerProProAsnAspIleSerThrSerValIleGlnSerGluValThrValProThrGlu 1134
DB 3901 AGCCCTCCCAATGACATTTCCACATCACTAATTCAGTGCAGGTACAGTCCACTGAA 3960
QY 1135 GluGlyIleAlaGlyValAlaThrSerThrGlyValAlaThrSerGlyLeuProIle 1154
DB 3961 GAGGGGAAATTCGTGGAGTTGCCACACAGACAGGTGTGGTACTTCAGGTGGTCTCCCA 4020
QY 1155 ProProValSerGluSerProValLeuSerSerValValSerSerIleThrIleProAla 1174
DB 4021 CCACCTGTGTGATCAACACAGTACTTCCAGCGTAGTTTCAAGTATCAACATACCTGCA 4080
QY 1175 ValValSerIleSerThrThrSerProSerIleGlnValProThrSerThrSerGluIle 1194
DB 4081 GTTGCTCAATATCTCACTACATCCCGTCACTTCAAGTCCCAACATCCACATCTGAGATC 4140
QY 1195 ValValSerSerThrAlaLeuThrProSerValThrValSerAlaThrSerAlaSerAla 1214
DB 4141 GTTGTTCTAGTACAGCACTGTATCTTCAATACAGTTTACAGCACTTCAGCTCTGCA 4200
QY 1215 GlyGlySerThrAlaThrProGlyProIysProProAlaValValSerGlnAlaAla 1234
DB 4201 GGGGGCACTACTGCTACCCACAGTCTTAAGCCTCCAGCTGATGATCTCAGACAGCACA 4260
QY 1235 GlySerThrThrValGlyAlaThrLeuThrSerValSerThrThrThrThrSerPheProSer 1254
DB 4261 GGGAGCACTACTGTGGAGGACACATTAACATCAGTTTACACACATCTCCCAATCCCAAGC 4320
QY 1255 ThrAlaSerGlnLeuSerIleGlnLeuSerSerSerThrSerThrProThrLeuAlaGlu 1274
DB 4321 ACAGCTTCACAGCTGTCACTTACAGTTAGCAGAGTACTTCTACCTCCACTTACGTGAA 4380
QY 1275 ThrValValValSerAlaHisSerLeuAspIysThrSerHisSerSerThrThrGlyLeu 1294
DB 4381 ACCGTGGATAGTAGGACACTCACTAGATTAAGATCTCATAGACAGTACAACTGGAATG 4440
QY 1295 AlaPheSerLeuSerAlaProSerSerSerSerProGlyValAlaGlyValSerSerTyr 1314
DB 4441 GCTTTTCCTCTCTGACCAATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4500
QY 1315 IleSerGlnProGlyGlyLeuHisProLeuValIleProSerValIleAlaSerThrPro 1334
DB 4501 ATTCTCAGCCTGGTGGGCTGCATCTTGTGATTCATATGATGATAGCTTCTTACTCT 4560
QY 1335 IleLeuProGlnAlaAlaGlyProThrSerThrProLeuLeuProGlnValProSerIle 1354
DB 4561 ATTCTTCCCAAGACAGACGACCTACTTCTACCTTATTAATCCCAAGTACTAGTATC 4620
QY 1355 ProProLeuValGlnProValAlaAsnValProAlaValGlnGlnThrLeuIleHisSer 1374
DB 4621 CCACCTTGTGATGAGCTGTGGCCAAATGTGCTGTGACAGCAGACCAATTAATCATAGT 4680
QY 1375 GlnProGlnProAlaLeuLeuProAsnGlnProHisThrHisCysProGluValAspSer 1394
DB 4681 CAGCCTCAACAGCTTGTGCTCCCAACAGCCCATACTCTGTCTCTGAAAGTATGATCT 4740
QY 1395 AsnThrGlnProIysAlaProGlyIleAspAspIleIysThrLeuGlnGluIysLeuArg 1414
DB 4741 GATACACAACCCAAAGCTCTCTGGAATGATGACATAAAGACTCTAGAAAGAAAGCTCGG 4800
QY 1415 SerLeuPheSerGluHisSerSerSerSerGlyAlaGlnHisAlaSerValSerLeuGluThr 1434
DB 4801 TCTCTGTCTGATGACACAGCTCAATCTGAGACTAGAGATGCTCTGTCTCACTGGAAGC 4860
QY 1435 SerLeuValIleGlnSerThrValThrProGlyIleProThrThrAlaValAlaProSer 1454
DB 4861 TCACCTAGTCAATAGAGCACTGTACACACAGGATCCCAACTACTGCTGTGGACCAAGC 4920
QY 1455 LysLeuLeuThrSerThrThrSerThrCysLeuProProThrAsnLeuProLeuGlyThr 1474
DB 4921 AAACCTCCGACTTTCACCAAGTACTCTGTACACCAACCAATTTACACTAGAGAA 4980
QY 1475 ValAlaLeuProValThrProValValThrProGlyGlnValSerThrProValSerThr 1494
DB 4981 GTTGCTTGCAGTATCACCAAGTGTGTCACCTGGGCAAGTTTCAACCCATCTGAGCACT 5040
QY 1495 ThrThrSerGlyValIysProGlyThrAlaProSerIysProProLeuThrIysAlaPro 1514
DB 5041 ACTACATAGAGAGTGAACCTGGAACTGCTCTCTCAAGCCCACTCACTAAGGCTCGG 5100
QY 1515 ValLeuProValGlyThrGluLeuProAlaGlyThrLeuProSerGluGlnLeuProPro 1534
DB 5101 GTGCTGCAGTGGGATACGAACTTCCACAGGTACTTACCCAGGACAGAGCTGCCACT 5160
QY 1535 PheProGlyProSerLeuThrGlnSerGlnProLeuGlnAspLeuAspAlaGlnLeu 1554
DB 5161 TTTCAGGACCTTCTCTAACCACAGTCCACGAACTCTAGAGATCTTGATGCTCAATTG 5220
QY 1555 ArgArgThrLeuSerProGlu***IleThrValThrSerAlaValGlyProValSerMet 1574
DB 5221 AGAAGAACACTTAGTCCAGATATATCAAGTACCTTCTGGGTTGCTGTGTCCATG 5280
QY 1575 AlaAlaProThrAlaIleThrGluAlaGlyThrGlnProGlnIysGlyValSerGlnVal 1594
DB 5281 GCGGCTCCAAACAGCAATACAGAAAGCAGAAACAGCTTCAGAAAGGTGTTCTCAAGTC 5340
QY 1595 LysGluGlyProValIleAlaThrSerSerGlyAlaGlyValPheIysMetGlyArgPhe 1614
DB 5341 AAAAGAGGCCCTGTCCATGCAACTAGTTCAGGAGCTGTGTTTAAAGATGGAGCATTT 5400
QY 1615 GlnValSerValAlaAlaAspGlyAlaGlnIysGluGlyIysAsnIysSerGluAspAla 1634
DB 5401 CAGGTTCTGTGGAGAGAGAGGAGTCCAGAAAGGGTAAATTAAGTCAAGAAATGCA 5460
QY 1635 LysSerValHisPheGlnSerSerThrSerGlnSerSerValLeuSerSerSerPro 1654
DB 5461 AAGCTGTTCATTTTGATCCAGACCTCAGAGTCTCAGTGCATATCAAGTATGATGCA 5520
QY 1655 GluSerThrLeuValIysProGlnProAsnGlyIleThrIleProGlyIleSerSerAsp 1674
DB 5521 GAGAGTACCTTGTGGAAACAGAGCCGAATGGCATTAACATCCCTGTATCTCTCAGAT 5580
QY 1675 ValProGlnSerAlaHisIlysThrThrAlaSerGlnAlaIysSerAspThrGlyGlnPro 1694
DB 5581 GTGCCAGAGGTGCCCAAAATCTAGCTCTCAGAGCAAAAGTACACATCGGACCT 5640
QY 1695 ThrIysValGlyArgPheGlnValThrThrThrAlaAsnIysValGlyArgPheSerVal 1714
DB 5641 ACCAAGTTGGAGCTTTTCAAGTGCACATACAGCAAAAGTGGTCTTTCTCTGTA 5700
QY 1715 SerIysThrGluAspIysIleThrAspThrIysIysGlnGlyProValAlaSerProPro 1734
DB 5701 TCAAAAATGAGGACAAAGATCAGTACACAAAGAAAGAAAGAACAGAGGATCTCTCT 5760
QY 1735 PheMetAspLeuGlnGlnAlaValLeuProAlaValIleProIysIysGluIysProGlu 1754
DB 5761 TTTATGATTTGGAACAAGCTGTCTTCTCTGCTGTGTATCCAAAGAAAGAAAGCTGAA 5820
QY 1755 LeuSerGlnProSerHisIleAsnGlyProSerSerAspProGlnAlaAlaPheLeuSer 1774
DB 5821 CTGTACAGAGCTTCACATCTAAATGGGCGCTTCTGACCCGAGAGGCGCTTTTAAAT 5880
QY 1775 ArgAspValAspAspGlySerGlySerProHisSerProHisGlnLeuSerSerIysSer 1794
DB 5881 AGGATGTGATATGATGTCCGTAGTCCACATCGGCCCATAGCTGAGCTCAAGAGC 5940
QY 1795 LeuProSerGlnAsnLeuSerGlnSerLeuSerAsnSerPheAsnSerTyrMetSer 1814
DB 5941 CTTCCTAGCCAGATATCTAATCAAGCTTAAAGCTTAAATTCATTAACTCTCTTCAATAGT 6000
QY 1815 SerAspAsnGluSerAspIleGluAspGluAspLeuIysLeuArgArgLeuArg 1834

Db	6001	AGCGACAAATGAGTCAGATATCGAAGATGAAAGACTTAAAGTTAAGGCTGGACAGACTTAACGA	6060
Qy	1835	AspLysHisLeuLysGluIleGlnAspLeuGlnSerArgGlnLysGluIleGluSer	1854
Db	6061	GATAAACATCTCAAAAGAGATTTCAGACCTCAGAGTGCAGACGAAGCATGAAATTGAATCT	6120
Qy	1855	LeuThrThrLysLeuGlyLysValProProAlaValIleIleProProAlaAlaProLeu	1874
Db	6121	TTGTATACCAAACTGGGCGAAGGTGCCCCCTGCTGTTATTATTTCCCGAGTGTCTCCCTT	6180
Qy	1875	SerGlyArgArgArgProThrLysSerLysGlySerLysSerSerArgSerSerSer	1894
Db	6181	TCAGGGAGAGACACACACCCTTAAGCAAAAGCAAGCAAAATCTAGTCCAGACAGATTC	6240
Qy	1895	LeuGlyAsnLysSerProGlnLeuSerGlyAsnLeuSerGlyGlnSerAlaAlaSerVal	1914
Db	6241	TTGGGGAATATAAGCCCCCAGCTTTCAGGTAACTGCTGTGTGATGAGATGCAGCTTCAGTC	6300
Qy	1915	LeuHisProGlnGlnThrLeuHisLeProProGlyAsnIleProGluSerGlyGlnIleGln	1934
Db	6301	TTGCACCCCCACGACAGACCTTCACCCCTCGGCAAACTCCAGATCCAGGCTCGGCGAATCAG	6360
Qy	1935	LeuLeuGlnProLeuLysProSerProSerSerAspAsnLeuThrSerAlaPheThrSer	1954
Db	6361	CTGTTACAGCCCCCTTAAGCAATCTCCCTCCAGTACAACTCTATTAGCTTCACACAGT	6420
Qy	1955	AspGlyAlaIleSerValProSerLeuSerAlaProGlyGlnGlyThrSerSerThrAsn	1974
Db	6421	GATGGTGCATTTTCAGTACCAAGCTTTCTGCTCAGGTCAAGGAAACAGAGACAAAC	6480
Qy	1975	ThrValGlyAlaThrValAsnSerGlnAlaAlaGlnAlaGlnProProAlaMetThrSer	1994
Db	6481	ACTGTTGGGGCAACAGTGAACAGCCAGCCCCCAAGCTCAGCTCTGCATTAACCTCC	6540
Qy	1995	SerArgLysGlyThrPheThrAspAsnLeuHisLysLysLeuValAspAsnTrpAlaArgAsp	2014
Db	6541	AGCAGGAAGGGGACATTCACAGATGACTTGACAAAGTTGGTACCAATTGGGCCCGAGAT	6600
Qy	2015	AlaMetAsnLeuSerGlyArgArgGlySerLysGlyHisMetAsnTyrGlnGlyProGly	2034
Db	6601	GCCATGATCTCTCAGCGACGAGGAGGAAGCAAGGAGCATGATTTAGAGGGCCCTGGA	6660
Qy	2035	MetAlaArgLysPheSerAlaProGlyGlnLeuCysIleSerMetThrSerAsnLeuGly	2054
Db	6661	ATGCGAAGGAAGTCTCTGCAACTGGGGCACTGTGCATCTCCAGACCTCGAACCCTGGGT	6720
Qy	2055	GlySerAlaProIleSerAlaAlaSerAlaThrSerLeuGlyHisPheThrLysSerMet	2074
Db	6721	GAGCTGTGCCCCCATCTTCGACAGCATCACTCTCTTAGGTGCACTTCACCAAGTCTATG	6780
Qy	2075	CysProProGlnGlnTyrGlyPheProAlaThrProPheGlyAlaGlnTrpSerGlyThr	2094
Db	6781	TGCCCCCAGACAGCACTATGAGCTTCCAGCTACCCATTGGGCGCTCAATGAGGGAGCG	6840
Qy	2095	GlyGlyProAlaProGlnProLeuGlyGlnPheGlnProValGlyThrAlaSerLeuGln	2114
Db	6841	GGTGGCCAGACACACAGCACTTGCGCAGTTCCAACTGTGGGAACCTGCTCTTGCGG	6900
Qy	2115	AsnPheAsnIleSerAsnLeuGlnLysSerIleSerAsnProProGlySerAsnLeuArg	2134
Db	6901	AATTTTCAAATCAGCAATTTGCAAAATTCATCAACCAACCCCAAGCTCCAACTCGCGG	6960
Qy	2135	ThrThr 2136	
Db	6961	ACCACT 6966	
RESULT 13			
ACH03794			
ID ACH03794 standard; DNA; 6147 BP.			
XX ACH03794;			
DT 26-SEP-2003 (first entry)			

XX	DNA encoding novel human protein kinase #31.
XX	Human; kinase; ds; gene; cosmetic application; nutraceutical application.
XX	Homo sapiens.
XX	US6541252-B1.
XX	01-APR-2003.
XX	14-MAY-2001; 2001US-00854856.
XX	19-MAY-2000; 2000US-0206015P.
XX	(LEXI-) LEXICON GENETICS INC.
PA	Walke DW, Halbur E, Donoho G, Turner CA;
PI	WPI; 2003-575927/54.
DR	F-PDB; ABO44411.
XX	New nucleic acid encoding novel human proteins, useful in cosmetic and
PT	nutraceutical applications.
PS	Disclosure; Page; 11pp; English.
XX	The invention relates to a new isolated nucleic acid encoding a novel
CC	human protein kinase. The nucleic acid is useful in cosmetic and
CC	nutraceutical applications. The present sequence represents DNA encoding
CC	a novel human protein kinase. Note: The sequence data for this patent did
CC	not from part of the printed specification but was obtained in electronic
CC	format directly from USFTO at
CC	segdata.uspto.gov/sequence.html?DocID=6541252B1
SQ	Sequence 6147 BP; 1727 A; 1630 C; 1393 G; 1395 T; 0 U; 2 Other;
Alignment Scores:	
Pred. No.:	0 Length: 6147
Score:	10344.00 Matches: 2048
Percent Similarity:	98.65% Conservative: 0
Best Local Similarity:	Mismatches: 0
Query Match:	95.67% Indels: 28
DB:	Gaps: 1
US-10-010-720-14 (1-2136) x ACH03794 (1-6147)	
QY	61 MetAaplySASPSeArgGlyAlaAlaThrThrThrThrGluHisArgPhe 80
Db	1 ATGGACAGGAACGCCGTGGGCGGC CGGACCATTACCAACACATGAGACGCTTCTTC 60
QY	81 ArgArGSerValIleCysAspSerAnaIthrAlaLeuGluLeuProGlyLeuProLeu 100
Db	61 CGCCGAGGCGTCATCTGCACATCCAAATGCACATGCATGACATGACTTCCGGCTTCTCTT 120
QY	101 SerLeuProGlnProSerIleProAlaAlaValProGlnSerAlaProProGluProHis 120
Db	121 TCCTTCGCCACGCCCATGCCCCCGGGCTGTCCGCAAGTGCTCACCGAGCCCTC 180
QY	121 ArgGluGluThrValThrAlaThrAlaThrSerGlnValAlaGlnGlnProProAla 140
Db	181 CGGAGAGAACCTGACCCGCCACGCCACTTCCAGGTAGCCAGAGCTCCAGCCGCT 240
QY	141 AlaAlaProGlyGluGlnAlaValAlaGlyProAlaProSerThrValProSerSerThr 160
Db	241 GCCGCCCTTGAGGAACAGGCCCTGCGGGCCCTGCCCCCTGCATCTCCCGACGTACC 300
QY	161 SerLyAspArgProValSerGlnProSerLeuValGlySerLyGlnGluProProPro 180
Db	301 AGCAAAAGACCGCCCAATGTCCAGCCTTGTGTGGGAGCAAGAAGACGCGCCGCG 360
QY	181 AlaArgSerLySerGlyGlySerAlaValGluProGlnGluGluArgSerGlnGln 200

```

Db      361 GCGAAGAGTGGACGCGCGCGCAGCCGCAAGGACCAAGAGGAAAGGACCGACGAC 420
QY      201 GlnAspAspIleGluGluLeuGluThrLysAlaValGlyMetSerAspAspGlyArgPhe 220
Db      421 CAGGATGATATCGAAGAGCTGGAGAACGAGCGGAGATGTCTPAACATGCGCGCTTT 480
QY      221 LeuLysPheAspIleGluIleGlyArgGlySerPheLysThrValTyrLysGlyLeuAsp 240
Db      481 CTCAAGTTTGACATCGAAATCGGACAGAGCTCTTTAAGAGGTCTACAAAGGTCTGGAC 540
QY      241 ThrGluThrThrValGluValAlaIleProCysGluLeuGlnAspArgLysLeuThrLysSer 260
Db      541 ACTGAACCAACCGGTGAAGTCCCGGTGTGTGAACGAGATCGAAATTAACAAGTCT 600
QY      261 GluArgGlnArgPheLysGluGluValIleGluMetLeuLysGlyLeuGlnHisProAsnIle 280
Db      601 GAGAGGACGAGATTTAAAGAGAGACCTGAATGTTAAAGGTCCTTCAGCATCCCAATAT 660
QY      281 ValArgPheTyrAspSerTyrPgluSerThrValLysGlyLysCysIleValLeuVal 300
Db      661 GTTAGATTTTATGATTCCTGGGAATCCACAGTAAAGGAAGAGTGTGTTTGGTG 720
QY      301 ThrGluLeuMetThrSerGlyThrLeuLysThrTyrLeuLysArgPheLysValMetLys 320
Db      721 ACTGACTTATGACGTCTGGAACTGGAACCTTAAACGTAATCGAAAGGTTTAAAGTGAATGAG 780
QY      321 IleLysValLeuArgSerTyrPgluSerGlnIleLeuLysGlyLeuGlnPheLeuHisThr 340
Db      781 ATCAAGTTCATAGAGAGTGGTGGCCGTGATCCTTAAAGGTCCTTCACATTTCTTCATACT 840
QY      341 ArgThrProProIleIleHisArgAspLeuLysCysAspAsnIlePheIleThrGlyPro 360
Db      841 CGAATCCACCTATCATTCACCGGATCTTAAATGTGCAACATCTTATACCGCGGCTT 900
QY      361 ThrGlySerValLysIleGlyAspLeuGlyLeuAlaThrLeuLysArgAlaSerPheAla 380
Db      901 ACTGGCTCAGTCAGATTTGGAGACCTCGGTCTGGCAACCTGAAACCGGCTCTCTTTGCC 960
QY      381 LysSerValIleGlyThrProGluPheMetAlaProGluMetTyrGluGluLysTyrAsp 400
Db      961 AAGAGTGTGATGATGACCCGAGAGTTTCAATGCCCCCTGAGATGTAGAGAAATATGAT 1020
QY      401 GluSerValAspValTyrAlaPheGlyMetCysMetLeuGluMetAlaThrSerGluTyr 420
Db      1021 GAAACCCCTTGACGTTTATGCTTTTGCGATGTGATGCTTGAGTGGCTACATCGAATAT 1080
QY      421 ProTyrSerGluCysGlnAsnAlaIleGlnIleTyrArgAlaGlyAlaThrSerGlyValLys 440
Db      1081 CTTTACTCGAGTGCAGAAATGTGCAACAGTCTAACCGTCCGCGACAGTGGGGTGAAG 1140
QY      441 ProAlaSerPheAspLysValAlaIleProGluValLysGluIleIleGluGlyCysIle 460
Db      1141 CCGAGCCAGTTTGGCAAGATGACATTCCTGAGATGAGAAATTTATTTGAAGATGACATA 1200
QY      461 ArgGlnAsnLysAspGluArgTyrSerTyrLysAspLeuAsnHisAlaPhePheGln 480
Db      1201 CGACAAAACAAAGATGAAAGATATTCATCAAGACCTTTGAACCATGCTTCTTCCA 1260
QY      481 GlnGluThrGlyValArgValGluLeuAlaGluGluAspAspGlyGlyLysIleAlaIle 500
Db      1261 GAGGAAACAGAGATGACGGGTGAATTTAGCAGAAAGACATGATGAGAAAAAATAGCCATA 1320
QY      501 LysLeuTyrPheLysArgIleGluAspIleLysLysLeuLysGlyLysTyrLysAspAsnGlu 520
Db      1321 AAATTTATGCTACGATATTAAGATATTTAAGAAATTTAAAGGAAATATCAAAAGATATGAA 1380
QY      521 AlaIleGluPheSerPheAspLeuGluArgAspValProGluAspValAlaGlnGluMet 540
Db      1381 GCATTTGAGTTTCTTTCTTTGATTAGAGAGATGTCCCAAGAGTGTGCAAGAAATG 1440
QY      541 ValGluSerGlyTyrValCysGluGlyAspHisLysThrMetAlaLysAlaIleLysAsp 560
Db      1441 GTTAGGTCTGGGTATGCTGTGAAGGTGATCAAGACATGGCTTAAAGCTTACAAAGAC 1500

```

```

QY      561 ArgValSerLeuIleLysArgLysArgGluGlnArgGlnLeuValArgGluGluGln 580
Db      1501 AAGATATCATTAATTAAGAGAAACGAGAGCAGCGGAGTTGTACGGGAGGACGAGAA 1560
QY      581 LysLysLysGlnGluGluSerSerLeuLysGlnValGluGlnSerSerAlaSerGln 600
Db      1561 AAAAAAACCGAGAGAGAGAGCTCTCAACAGCAGGTAGAAACAATCCAGTCTTCCGAG 1620
QY      601 ThrGlyIleLysGlnLeuProSerAlaSerThrGlyIleProThrAlaSerThrThrSer 620
Db      1621 ACAGGAATCAAGACGCTCCCTTCTGCTAGCACCGGCATACCTACTGCTTCTACCACTTCA 1680
QY      621 AlaSerValSerThrGlnValGluProGluGluProGluAlaAspGlnHisGlnGlnLeu 640
Db      1681 GCTTCAGTTTTCACACAGTGAACCTGAAACCTGAGGACGATCAACATCAACACTA 1740
QY      641 GlnTyrGlnGlnProSerIleSerValLeuSerAspGlyThrValAspSerGlyGlnGly 660
Db      1741 CAGTACGACAAACCGATATCTGTGTATCTGATGAGAGCGTTGACAGTGTCTCAGGGA 1800
QY      661 SerSerValPheThrGluSerArgValSerSerGlnGlnThrValSerTyrGlySerGln 680
Db      1801 TCCTCTGCTTCAAGAAATCTCGAGTAGCAGCAACAGACAGTTTCATATGTTCCCA 1860
QY      681 HisGluGlnAlaHisSerThrGlyThrValProGlyHisIleProSerThrValGlnAla 700
Db      1861 CATAAACAGGACACATTTTCACAGGACAGTCCAGGGCATATATCTTCACTGATCCCAACA 1920
QY      701 GlnSerGlnProHisGlyValTyrProProSerSerValAlaGlnGlyGlnSerGlnGly 720
Db      1921 CAGTCTCAGCCCATGCGGTATATCCACCTCAAGTGTG----- 1959
QY      721 GlnProSerSerSerSerLeuThrGlyValSerSerSerGlnProIleGlnHisProGln 740
Db      1959 ----- 1959
QY      741 GlnGlnGlnGlnGlyIleGlnGlnThrAlaProProGlnGlnThrValGlnTyrSerLeuSer 760
Db      1960 ---CAGAGGAAATACAGCAGACGCCCCCTTCCAAACAGACAGGAGATATTCACCTTCA 2016
QY      761 GlnThrSerThrSerSerGlnAlaThrThrAlaGlnProValSerGlnProGlnAlaPro 780
Db      2017 CAGACATCAACCTCAGATGAGGCCACTACTGCAACGCACTGAGTCAAGCTTCAAGCTTCA 2076
QY      781 GlnValLeuProGlnValSerAlaGlyLysGlnSerThrGlnGlyValSerGlnValAla 800
Db      2077 CAAGTCTTGCTCAAGTATCAGCTGGAACAAAGATCTCAGGAGTCTCAGGTTGCT 2136
QY      801 ProAlaGluProValAlaValAlaGlnProGlnAlaThrGlnProThrThrLeuAlaSer 820
Db      2137 CTTGCAAGAGCCAGTGTGAGTAGCAACGCCCAAGCTAACCCAGCCGACACACTTGTGCTTCC 2196
QY      821 SerValAspSerAlaHisSerAspValAlaSerGlyMetSerAspGlyAsnGluAsnVal 840
Db      2197 TCTGTAGACAGTGCACATTACAGATGTTGCTTCAAGTATGATGATGCAATGAGAAAGTTC 2256
QY      841 ProSerSerSerGlyArgHisGluGlyArgThrThrLysArgHisTyrArgLysSerVal 860
Db      2257 CCATCTTCCAGTGAAGGACATGAAGAGAACTACAAACAGGCACTTACCAAAATCTGTA 2316
QY      861 ArgSerArgSerArgHisGlyLysThrSerArgProLysLeuArgIleLeuAsnValSer 880
Db      2317 AGGAGTGGCTCTGACATGAAAAAAGCTTCAGCCCAAAATTAAGAAATTTTGAAATGTTTCA 2376
QY      881 AsnLysGlyAspArgValValGluCysGlnLeuGlnThrHisAsnArgLysMetValThr 900
Db      2377 AATTAAGAGACCGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2436
QY      901 PheLysPheAspLeuAspGlyAspAsnProGluGluIleAlaThrIleMetValAsnAsn 920
Db      2437 TTCAAATTTGACCTAGATGTGACAAACCCGAGAGATGACAAATTAATGATGATGATGATGAT 2496

```


QY 921 AspPheIleLeuAlaIleGluArgGluSerPheValAspGlnValArgGlnIleIleGlu 940
Db 2497 GACCTTATCTCAGCAATAGAGAGAGAGTCCGTTTGATCAAGGCAAAATTATGAA 2556
QY 941 LysAlaAspGluMetLeuSerGluAspValSerValGluProGluGluAspGlnGluLeu 960
Db 2557 AAAGCGATGAATGCTCACTGAGAGATGTCAGTGTGAAACCAAGGGGTGATCGGGGATG 2616
QY 961 GluSerLeuGlnGluLysAspAspTyrGluPheSerGlySerGlnLysLeuGluGlu 980
Db 2617 GAGAGCTTACAGAGAAAGAGATGACTATGCTTTTCAAGTCTCAGAAATTGGAAGAGAG 2676
QY 981 PheLysGlnProIleProAlaSerSerMetProGlnGlnIleGlyIleProThrSerSer 1000
Db 2677 TTCAACCAACCAATCTCGGCGTCTTCATGCAACAGCAAAATAGGCATTTCTACAGTTCT 2736
QY 1001 LeuThrGlnValValHisSerAlaGlyValArgArgPheIleValSerProValProGluSer 1020
Db 2737 TTAACTCAAGTTGTTCTTCTGCGGAAAGCGGTTTATGAGTCTCTGTCGCAAGAGC 2796
QY 1021 ArgLeuArgGluSerLysValPheProSerGlnIleThrAspThrValAlaAlaSerThr 1040
Db 2797 CGATTAGAGAGATCAAAAGTTTCCCAAGTGAATTAACAGATACAGTTGCTGCTCTACA 2856
QY 1041 AlaGlnSerProGlyMetAsnLeuSerHisSerAlaSerSerLeuSerLeuGlnAla 1060
Db 2857 GCTCAGAGCCCTGGAATGAAGTCTGCTCAGTCTCATCATCTAGTCTGTCACAGAGCC 2916
QY 1061 PheSerGluLeuValArgValGlnMetThrGluGlyProAsnThrAlaProProAsnPhe 1080
Db 2917 TTTTCTGAACTTACAGCGTCCCAATGACAGAAAGAGCCCAAVACAGACCTCCCAAACTTT 2976
QY 1081 SerHisThrGlyProThrPheProValValProPheLeuSerSerIleAlaGlyVal 1100
Db 2977 AGTATACAGAGCAACATTTCCAGTAGTACCTCTCTTTTATGATGATTCCTGAGATC 3036
QY 1101 ProThrThrAlaAlaAlaThrAlaProValProAlaThrSerSerProProAsnAspIle 1120
Db 3037 CCAACCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3096
QY 1121 SerThrSerValIleGlnSerGlnValThrValProThrGluGluGlyIleAlaGlyVal 1140
Db 3097 TCCACATCAGTAAATTCAGTCTGAGGTACAGTGCCTGAAAGGGGATTCGCGAGTT 3156
QY 1141 AlaThrSerThrGlyValValThrSerGlyLeuProIleProProValSerGlnSer 1160
Db 3157 GCCACCAAGAGAGGTGAGTCACTCAGTGTGCTCCCATACCACTGTGTCTGATCA 3216
QY 1161 ProValLeuSerSerValValSerSerIleThrIleProAlaValValSerIleSerThr 1180
Db 3217 CCAGTACTTCCAGGTAGTTCAGATACAGTCAAAATCACTGCAAGTTGCTCAATATCTACT 3276
QY 1181 ThrSerProSerLeuGlnValProThrSerThrSerGluIleValValSerSerThrAla 1200
Db 3277 ACATCCCGGTCACTCAAGTCCCAATCAATCAATCAATCAATCAATCAATCAATCAATCA 3336
QY 1201 LeuThrProSerValThrValSerAlaThrSerAlaSerAlaGlyIleSerThrAlaThr 1220
Db 3337 CTGATCTCTCAGTACAGTTCAGCACTTCAAGCTCTGAGGGGGGAGTACTGCTACC 3396
QY 1221 ProGlyProLysProProAlaValValSerGlnAlaAlaGlySerThrThrValGly 1240
Db 3397 CCAAGTCTTACAGCTCCAGCTGATATCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3456
QY 1241 AlaThrLeuThrSerValSerThrThrThrSerPheProSerThrAlaSerGlnLeuSer 1260
Db 3457 GCCACATTAACATCAGTTTCTACCACTCATTCCTCAAGAGAGAGAGAGAGAGAGAGAG 3516
QY 1261 IleGlnLeuSerSerSerThrSerThrProThrLeuAlaGluThrValValValSerAla 1280
Db 3517 ATTACGCTTACAGAGAGTACTTCACTCTTACCTTTAGCTGAACCGGTGATGAGCGCA 3576
QY 1281 HisSerLeuAspLysThrSerHisSerThrThrGlyLeuAlaPheSerLeuSerAla 1300

Db 3577 CACTCACTGATAGAGCATCTTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3636
QY 1301 ProSerSerSerSerSerProGlyAlaGlyValSerSerTyrIleSerGlnProGlyGly 1320
Db 3637 CCATCTTCTCTCTCTCTCTCTGAGAGAGAGAGTCTATATTTTCTAGCCGTGGGG 3696
QY 1321 LeuHisProLeuValIleProSerValIleAlaSerThrProIleLeuProGlnAlaAla 1340
Db 3697 CTGCATCTCTTGGTCACTTCCATGATGATGATCTTCACTCTTCTTCTTCCCAAGAGCA 3756
QY 1341 GlyProThrSerThrProLeuLeuProGlnValProSerIleProProLeuValGlnPro 1360
Db 3757 GGAAGTCACTTCAAGCTTATTAATCCCAAGTCAAGTATCCCAAGCTTGTGAGAGAGCT 3816
QY 1361 ValAlaAsnValProAlaValGlnGlnThrLeuIleHisSerGlnProGlnProAlaLeu 1380
Db 3817 GTTGCCATGTCCTGCTGTCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3876
QY 1381 LeuProAsnGlnProHisThrHisCysProGluValAspSerAspThrGlnProValAla 1400
Db 3877 CTTCACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3936
QY 1401 ProGlyIleAspAspIleIleThrLeuGlnLysLeuArgSerLeuPheSerGlnHis 1420
Db 3937 CCTGGAATGATACATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3996
QY 1421 SerSerSerGlyAlaGlnHisAlaSerValSerLeuGluThrSerLeuValIleGluSer 1440
Db 3997 AGCTCATCTGAGAGTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4056
QY 1441 ThrValThrProGlyIleProThrThrAlaValAlaProSerLysLeuLeuThrSerThr 1460
Db 4057 ACTGTACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4116
QY 1461 ThrSerThrCysLeuProProThrAsnLeuProLeuGlyThrValAlaLeuProValThr 1480
Db 4117 ACAGTACTGCTTACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4176
QY 1481 ProValValThrProGlyGlnValSerThrProValSerThrThrThrSerGlyValLys 1500
Db 4177 CCAAGTGTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4236
QY 1501 ProGlyThrAlaProSerLysProProLeuThrLysAlaProValLeuProValGlyThr 1520
Db 4237 CCGTGAAGTCTCTCTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4296
QY 1521 GluLeuProAlaGlyThrLeuProSerGlnLeuProProPheProGlyProSerLeu 1540
Db 4297 GAAGTCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4356
QY 1541 ThrGlnSerGlnGlnProLeuGluAspLeuAspAlaGlnLeuArgThrLeuSerPro 1560
Db 4357 ACCAGTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4416
QY 1561 Glu**IleThrValThrSerAlaValGlyProValSerMetAlaAlaProThrAlaIle 1580
Db 4417 GAGATKATCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4476
QY 1581 ThrGluAlaGlyThrGlnProGlnLysGlyValSerGlnValLysGluGlyProValLeu 1600
Db 4477 ACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4536
QY 1601 AlaThrSerSerGlyAlaGlyValPheLysMetGlyArgPheGlnValSerValAlaAla 1620
Db 4537 GCACTAGTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4596
QY 1621 AspGlyAlaGlnLysGluGlyLysAsnLysSerGluAspAlaLysSerValHisPheGlu 1640
Db 4597 GACGGTCCCAAGAAAGAGGTAAATAATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4656
QY 1641 SerSerThrSerGluSerSerValLeuSerSerSerProGluSerThrLeuValLys 1660

```

Db      4657  TCACGACCTCAGAGTCTCAGTGCATCAAGTAGTCCAGAGAGTACTTGATGAAA 4716
QY      1661  ProGluProAsnGlyIleThrIleProGlyIleSerSerAspValProGluSerAlaHis 1680
Db      4717  CCAGAGCCGAAATGCGATTAACATCCCTGATCTCTCAGATGTCGCAAGAGTCCAC 4776
QY      1681  LysThrThrAlaSerGlyAlaLysSerAspThrGlyGlnProThrLysValGlyThrPhe 1700
Db      4777  AAAAATACTAGCTCAGAGGCAAAATGACACACTGGGAGCCCTACCAAGTTGAGCTTT 4836
QY      1701  GlnValThrThrThrAlaAsnLysValGlyArgPheSerValSerLysThrGluAspLys 1720
Db      4837  CAGGTGACAACTACAGCAAAACAAAGTGGTCTCTCTCTGATCAAAAATCGAGAGCAAG 4896
QY      1721  IleThrAspThrLysLysGlyLysProValAlaSerProProPheMetAspLeuGln 1740
Db      4897  ATCATGACACAAAGAAAGAGACCAAGTGCATCTCTCTCTTATGATTTGGAACAA 4956
QY      1741  AlaValLeuProAlaValIleProLysLysGlyLysProGluLeuSerGluProSerHis 1760
Db      4957  GCTGTTCTCTCTGCTGATACCAAGAAAGAGAGCTGAACTGTCAAGGCTTCAAT 5016
QY      1761  LeuAsnGlyProSerSerSerProGluAlaAlaPheLeuSerArgAspValAspAspGly 1780
Db      5017  CTAAATGGGCGCTCTTGTGACCCGAGGCGCTTTTAAAGTAGGATGTGATGATGAT 5076
QY      1781  SerGlySerProHisSerProHisGlnLeuSerSerLysSerLeuProSerGluAsnLeu 1800
Db      5077  TCCGGTACTCCACACTGCCCATCAGCTGAGCTAAAGAGCTTCTTCAAGCAGATCTA 5136
QY      1801  SerGlnSerLeuSerAsnSerPheAsnSerSerLysSerLysSerSerAspAsnGlySer 1820
Db      5137  AGTCAAGACCTTAGTAATTCATTAACTTAACTCTTACATGAGTAGGAGCAATGAGTCAT 5196
QY      1821  IleGluAspGlyLysLeuLeuLeuGlnLeuArgArgLeuAlaAspLysHisLeuLysGlu 1840
Db      5197  ATCGAAGATGAAGCTTAAAGTTAGAGCTGCGACGACCTAAACATCTCAAAAGAG 5256
QY      1841  IleGluAspLeuGlnSerArgGlnLysHisGlnIleGlnSerLeuLysThrLysLeuGly 1860
Db      5257  ATTCAAGACCTCGAGAGTCCGCGCAGAGATCAAAATTGAAATTTTGATATCCAAACTGGGC 5316
QY      1861  LysValProProAlaValIleIleProProAlaAlaProLeuSerGlyArgArgArgArg 1880
Db      5317  AAGGTGCCCCCTGCTTATTAATCCCCCAAGCTCTCCCTTTCAGGAGAGAGAGAGAG 5376
QY      1881  ProThrLysSerLysGlySerLysSerSerArgSerSerSerLeuGlyAsnLysSerPro 1900
Db      5377  CCCACTAAAGCAAAAGGACGAAATCTAGTCGAGAGCTCTCTGGGGAATAAAGCCCC 5436
QY      1901  GlnLeuSerGlyAsnLeuSerGlyGlnSerAlaAlaSerValLeuHisProGlnGlnThr 1920
Db      5437  CAGCTTTCAGGTAACTGCTGCTGCTGACAGTCCAGCTTCAAGCTTTCGACCCCCAGAGACC 5496
QY      1921  LeuHisProProGlyAsnIleProGlnSerGlyGlnAsnGlnLeuGlnProLeuLys 1940
Db      5497  CTCACACCTCTGCGCAACATCCAGAGTCCGGGAGAGATCAGCTTACAGCCCCCTTAAG 5556
QY      1941  ProSerProSerSerAspAsnLeuLysSerAlaPheThrSerAspGlyAlaIleSerVal 1960
Db      5557  CCATCTCTCTCAGTGAACCTTATTCAGCTTCAACGATGATGATGATGATGATGATGAT 5616
QY      1961  ProSerLeuSerAlaProGlyGlnGlyThrSerSerThrAsnThrValGlyAlaThrVal 1980
Db      5617  CCAAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5676
QY      1981  AsnSerGlnAlaAlaGlnAlaGlnProProAlaMetThrSerSerArgLysGlyThrPhe 2000
Db      5677  AACAGGCCAAGCCGCCAAGCTCAGCTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5736
QY      2001  ThrAspAspLeuHisLysLeuValAspAsnThrAlaArgAspAlaMetAsnLeuSerGly 2020
Db      5737  ACGATGACTTGCACAAAGTTGGTAGCAATTTGGGCCGAGATGCATGATCTCTCAGGC 5796

```

```

QY      2021  ArgArgGlySerLysGlyHisMetAsnTyrgLugIyProGlyMetAlaArgLysPheSer 2040
Db      5797  AGGAGAGAGAGAGCAAGAGGACATGATTAATGAGGGCCCTCGAATGGCAAGAAAGTTCTCT 5856
QY      2041  AlaProGlyGlnLeuGlyIleSerMetThrSerAsnLeuGlyLysSerAlaProIleSer 2060
Db      5857  GCACTGGGACAACTGTCATCTCCATGACCTGAACTGGGAGGCTGCCCCCAATCTCT 5916
QY      2061  AlaAlaSerAlaThrSerLeuGlyHisPheThrLysSerMetCysProProGlnGlnTy 2080
Db      5917  GCAGATCAGCTACCTCTCTAGTCACTTCAACAGATGATGATGATGATGATGATGATGAT 5976
QY      2081  GlyPheProAlaThrProPheGlyAlaIleGlnTrpSerGlyThrGlyProAlaProGln 2100
Db      5977  GCTTTCAGCTACCCCAATTTGGCGCTCAATGAGTGGGACGGGTGGCCACACACAG 6036
QY      2101  ProLeuGlyGlnPheGlnProValGlyThrAlaSerLeuGlnAsnPheAsnIleSerAsn 2120
Db      6037  CCATTTGGCCAGTTCCAACTGTGGAACTGCTCTTGCAGAAATTTCAACATCTCAGCAAT 6096
QY      2121  LeuGlnLysSerLysSerAsnProProGlySerAsnLeuArgThrThr 2136
Db      6097  TTGCAGAAATCCATCAGCAACCCCCAGGCTCCAACTGGGAGCCACT 6144

RESULT 14
ACH03791
ID      ACH03791 standard; DNA; 6426 BP.
XX      XX
AC      ACH03791;
XX      XX
DT      26-SEP-2003 (first entry)
XX      XX
DE      DNA encoding novel human protein kinase #28.
XX      XX
KW      Human; kinase; ds; gene; cosmetic application; nutraceutical application.
XX      XX
OS      Homo sapiens.
XX      XX
EN      US6541252-B1.
XX      XX
PD      01-APR-2003.
XX      XX
PF      14-MAY-2001; 2001US-00854856.
XX      XX
PR      19-MAY-2000; 2000US-0206015P.
XX      XX
PA      (LEXI-) LEXICON GENETICS INC.
XX      XX
PI      Walke DW, Hilbun E, Donoho G, Turner CA;
XX      XX
DR      WPI; 2003-575927/54.
XX      XX
DR      P-PSDB; ABO44408.
XX      XX
PT      New nucleic acid encoding novel human proteins, useful in cosmetic and
PT      nutraceutical applications.
XX      XX
PS      Disclosure; Page: 11pp; English.
XX      XX
CC      The invention relates to a new isolated nucleic acid encoding a novel
CC      human protein kinase. The nucleic acid is useful in cosmetic and
CC      nutraceutical applications. The present sequence represents DNA encoding
CC      a novel human protein kinase. Note: The sequence data for this patent did
CC      not from part of the printed specification but was obtained in electronic
CC      format directly from USPTO at
CC      seqdata.uspto.gov/sequence.html?DocID=6541252B1
XX      XX
SQ      Sequence 6426 BP; 1789 A; 1714 C; 1455 G; 1465 T; 0 U; 3 Other;

Alignment Scores:
Pred. No.: 0 Length: 6426
Score: 10287.50 Matches: 2048
Percent Similarity: 94.42% Conservative: 0

```

Best Local Similarity: 94.42%
 Query Match: 95.15%
 DB: 8
 Mismatches: 0
 Indels: 121
 Gaps: 2

US-10-010-720-14 (1-2136) x ACH03791 (1-6426)

QY 61 MetAspIysAspSerArgIylAlaAlaAlaThrThrThrThrGluHisArgPhePhe 80
 DB 1 ATGGACAAAGACACCGCTGGGGGGCGCCGACACCTACCAACCTAGACACCGCTTTC 60
 QY 81 ArgArgSerValIleCysAspSerAsnAlaThrAlaLeuGluLeuProGlyLeuProLeu 100
 DB 61 CGCGGAGCGCTCATCTCGACTCCCAATGCCACCTGCACTGAGAGCTTCCCGGCTTCTT 120
 QY 101 SerLeuProGluProSerIleProAlaAlaValProGluSerAlaProProGluProHis 120
 DB 121 TCCCTGCCACGCGCACATCCCGCGGCTGTCGCGAGAGTGTCTCCACCGAGACCCAC 180
 QY 121 ArgGluGluThrValThrAlaThrAlaThrSerGluValAlaGluGluProProAlaAla 140
 DB 181 CGGGAAGACACCGTGAACCGCCACCGCACTTCCAGGTAGCCGACGCGCTCCAGCGCT 240
 QY 141 AlaAlaProGlyGluGluAlaValAlaGlyProAlaProSerThrValProSerSerThr 160
 DB 241 GCCGCCCCCTGGGGAACAGCGCGTCCGCGGCGCTCCCTCGACTGTCTCCAGCAATACC 300
 QY 161 SerIlyAspArgProValSerGluProSerLeuValGlySerIlyGluGluProProPro 180
 DB 301 AGCAAAACCGCCACAGTGTCCAGCTTACCTTGTGGGACCAAGAGAGACCCCGCG 360
 QY 181 AlaArgSerGlySerGlyGlyGlySerAlaGlyGluProGluGluGluArgSerGluGlu 200
 DB 361 GCGGAAGTGGCAACGCGCGCGCGGACCGCCACAGGACCAAGAGAGAGAGAGAGAGAG 420
 QY 201 GluAspAspIleGluGluLeuGluThrIlyAlaValGlyMetSerAsnAspGlyArgPhe 220
 DB 421 CAGATGATATCGAAGAGCTGAGACCAAGCGCTGGGAATGTCTAAGATGAGCGCGCTT 480
 QY 221 LeuIlyPheAspIleGluIleGlyArgGlySerPheIlyThrValIlyIlyGlyLeuAsp 240
 DB 481 CTCAAGTTGACATCGAATGCGCAGAGGCTCTTAAAGCGTCTACAAAGAGTGGAC 540
 QY 241 ThrGluThrThrValGluValAlaAlaTrpCysGluLeuGluAspArgIlySerThrIlySer 260
 DB 541 ACTGAACACACCGTGGAGAGTGGCTCGTGTGATCGACGAGATCGAAATTAACAAGCT 600
 QY 261 GluArgGluArgPheIlyGluGluAlaGluMetLeuIlyGlyLeuGluHisProAsnIle 280
 DB 601 GAGAGCGACAGATTAAAGAGAGCTGAAATGTTAAAGGCTTCCACATCCCAATATT 660
 QY 281 ValArgPheIlyArgSerIlyGluSerThrValIlyGlyIlyIlyCysIleValLeuVal 300
 DB 661 GTTAATTTTATGATTTCTCGGAATCCACAGTAAAGAGAGAGAGAGAGAGAGAGAG 720
 QY 301 ThrGluLeuMetThrSerGlyThrLeuIlyThrIlyLeuIlySerPheIlyValMetIly 320
 DB 721 ACTGAACCTTAAGACCTCTGGACACTTAAAGTATCTGAAAGAGTTTAAAGTATGAG 780
 QY 321 IleIlyValLeuArgSerIlyCysArgGluIleLeuIlyGlyLeuGluPheLeuHisThr 340
 DB 781 ATCAAGTCTTAAGAGAGTGGTGGCGCTGAGATCTTAAAGGCTTCCAGTCTTCTATCT 840
 QY 341 ArgThrProProIleIleHisArgAspLeuIlyCysAspAsnIlePheIleThrGlyPro 360
 DB 841 CGAATCTCACCTATATTCACCGGATCTTAAATGACACATCTTATTCACCGGCTT 900
 QY 361 ThrGlySerValIlyIleGlyAspLeuGlyLeuAlaThrLeuIlyArgIlyAspPheAla 380
 DB 901 ACTGCTCATGTAAGATTGAGACCTCGGTCTGGACACCTGAAAGGCGCTTCTTTGCC 960
 QY 381 IlySerSerValIleGlyThrProGluPheMetAlaProGluMetIlyGluIlyIlyIlyAsp 400
 DB 961 AAGAGTGTATGATGATACCCAGAGTTCATGGCCCTCGAGATGTATGAGAGAAATATGAT 1020

QY 401 GluSerValAspValIlyAlaPheGlyMetCysMetLeuGluMetAlaThrSerGluIly 420
 DB 1021 GAATCCGTGACGTTATATGCTTTGGATGTGATGCTTGAATGCTACATCTGAAATAT 1080
 QY 421 ProIlySerGluCysGluSerAlaAlaGluIleIleIlyArgArgValThrSerGlyValIly 440
 DB 1081 CCTTACTCGGAGGCCAAATGCTGCACAGANTCTACCTGCTGACCACTGGGGTGAAG 1140
 QY 441 ProAlaSerPheAspIlyValAlaIleProGluValIlyGluIleIleGluGlyCysIle 460
 DB 1141 CCAGCCAGTTTACAAAGTAGCAATTCCTGAAGTAGAAGAAATTTATGAAGATGCATA 1200
 QY 461 ArgGluAsnIlyAspGluArgIlySerIleIlyAspLeuLeuHisAlaPhePheGlu 480
 DB 1201 CGACAAACCAAGATGAATATTCATCAAGACCTTTGAACCATGCTTCTTCCAA 1260
 QY 481 GluGluThrGlyValArgValGluLeuAlaGluGluAspGlyGlyIlyIleAlaIle 500
 DB 1261 GAGGAAACAGAGGTACCGGTAGAAATTAGCAAGAAAGATAGTAGAAGAAATTAAGCATA 1320
 QY 501 IlyLeuThrPheLeuArgIleGluAspIleIlyLeuLeuIlyGlyIlyIlyAspAsnGlu 520
 DB 1321 AAATTATGCTTACGATTTAGATTAAGATTTAAGAAATTAAGGAAATTAAGATTAAGCA 1380
 QY 521 AlaIleGluPheSerPheAspLeuGluArgAspValProGluAspValAlaGluGluMet 540
 DB 1381 GCTATGATGTTTCTTTGATTTAGAGAGAGATCCCAAGAGATGTTGCAACAGAAATG 1440
 QY 541 ValGluSerGlyIlyValIlyCysGluGlyAspHisIlySerMetAlaIlyAlaIleAsp 560
 DB 1441 GTAGAGTCTGGATATGCTGTGAAGATGATCACAAAGCCATGCTTAAAGCTATTAAGAC 1500
 QY 561 ArgValSerLeuIleIlySerArgIlyGluGluGluGluValArgGluGluGluGlu 580
 DB 1501 AGATATCATTTAATTAAGAGAAACGAGACGCGGAGTGTGTACGGAAGACCAAGAA 1560
 QY 581 IlyIlyIlyGluGluGluGluSerSerLeuIlyGluGluValGluGluSerAlaSerGlu 600
 DB 1561 AAAAAAAGCAGAGAGAGAGAGAGCTCAACAGCAGGTAGAACATCACTCAGCTTCCAG 1620
 QY 601 ThrGlyIleIlyGluGluProSerAlaSerThrGlyIleProThrAlaSerThrThrSer 620
 DB 1621 ACAGAAATCAAGAGCTCCCTCTGTCTAGACCGGACATCACTTCTTACCACTTCA 1680
 QY 621 AlaSerValSerThrGluValGluProGluGluProGluAlaAspGluHisGluGluLeu 640
 DB 1681 GCTTCAGTTTCTACCAAGTGAACCTGAAGACCTGAGCAGATCAACATCAACCACTA 1740
 QY 641 GluIlyGluGluProSerIleSerValLeuSerAspGlyThrValAspSerGlyGluGly 660
 DB 1741 CAGTACGACCAACCAATATCTGTGTTATCTGATGAGGAGGCTTGAACAGTGTCAAGGA 1800
 QY 661 SerSerValPheThrGluSerArgValSerGluGluGluThrValSerIlySerGlu 680
 DB 1801 TCCCTGCTTCAAGATCTCGAGTGAAGCCCAAGACAGAGTTTCTAATAGTCTCCCA 1860
 QY 681 HisGluGluAlaHisSerThrGlyThrValProGlyHisIleProSerThrValGluAla 700
 DB 1861 CATGAACGCAATTTACACGACAGCTCCAGGCAATATCACTTACTGATCCAGCA 1920
 QY 701 GluSerGluProHisGlyValIlyProProSerSerValAlaGluGlyGluSerGluGly 720
 DB 1921 CAGTCTCAGCCCAAGGGGTATTCACCTTCACCTCAAGTGG----- 1959
 QY 721 GluProSerSerSerLeuThrGlyAlaSerSerSerGluProIleGluHisProGlu 740
 DB 1959 ----- 1959
 QY 741 GluGluGluGlyIleGluGluThrAlaProProGluGluGluThrValGluIlySerLeuSer 760
 DB 1960 ---CAGCAGGGAATACAGCAGACGCGCTCTCTCAAGACAGTCAAGATATCACTTCA 2016

```

QY 761 GlnThrSerThrsSerSerGlnAlaThrThrAlaGlnProValSerGlnProGlnAlaPro 780
Db 2017 CAGACATCAACCTCCAGTAGGCGCACTACTGCACAGCCAGTAGTACAGCTCAAGCTCCA 2076
QY 781 GlnValLeuProGlnValSerAlaGlyLysGln----- 791
Db 2077 CAAAGTCCTGCTCAAGATCAAGCTGGAAAAAGGGCTTCCACCTGCAGCTGCACACAG 2136
QY 791 ----- 791
Db 2137 TACCCAGAGGATTCAAAATATGCTCCCTCTTCCAAAGTGCTTCTGTTGCATTCATTC 2196
QY 791 ----- 791
Db 2197 ACAAGTCTAACCCCTCCCATGCGCAGAGAAAGTACGGCTACACCTGGGACTTTCACCA 2256
QY 791 ----- 791
Db 2257 GTGGTGACGCTTATGTGAAATCAAACTTTTAGTTCTATGGGTGTGTAGAGAGAGAG 2316
QY 791 ----- 791
Db 2317 GTTCAAGTGTCCAGCCAGAGAGAGGAGTTAGACAAAGCCCACTAATCTCCACAGCA 2376
QY 792 -----SerThrGlnGlyValSerGlnValAlaProAlaGlnProValAlaVal 807
Db 2377 GCAGTTTTGAGAGAGTACTCAGGAGAGTCTCAGAGTTGCTCTCGCAGAGCCAGTTGCACTA 2436
QY 808 AlaGlnProGlnAlaThrGlnProThrThrLeuAlaSerSerValAspSerAlaHisSer 827
Db 2437 GCACAGCCCAAGGCTACCCAGCCAGCACTTGGCTTCTCTGTAAGACAGTGCACATTCA 2496
QY 828 AspValAlaSerGlyMetSerAspGlyAsnGlnAsnValProSerSerSerGlyAlaHis 847
Db 2497 GATGTGCTTGAAGTATGAGTATGATGCGCAATGACATCGCTCCAGTGGAGAGGAT 2556
QY 848 GlnGlnValThrThrLysArgHisLysArgLysSerValArgSerArgSerArgHisGln 867
Db 2557 GAAGAAAGAACTACAAACCGCATTTACGAAATCTGTAAAGAGTCGCTCGACATGAA 2616
QY 868 LysThrSerArgProLysLeuArgLysLeuAsnValSerAsnLysGlyAspArgValVal 887
Db 2617 AAAATCTCAGCCCAAAATTTAGAAATTTGATGTTCTCAATAAAGAGACCGGATGTA 2676
QY 888 GlnCysGlnLeuGlnThrHisAsnArgLysMetValThrPheLysPheAspLeuAspGly 907
Db 2677 GAATGTCAATTAGAGACTCAATTAATGAAATGTTTACATTCAATTTGACCTTAGATGT 2736
QY 908 AspAsnProGlnGlnLysLeuAlaThrLysMetValAsnAsnAspPheLeuAlaLysGln 927
Db 2737 GACAAACCCCAAGAGATAGCAACAATATGATGTAACATGACTTATTTCTAGCAATAGAG 2796
QY 928 ArgGlnSerPheValAspGlnValArgGlnLysLeuLysAlaAspGlnMetLeuSer 947
Db 2797 AGAGAGTCGTTTGTGATCAGATGCGAGAAATTAATTAAGAAAGCTGATGAAGCTCAGT 2856
QY 948 GlnAspValSerValGlnProGlnGlyAspGlnGlyLeuGlnSerLeuGlnGlyLysAsp 967
Db 2857 GAGAGATCTAGTGTGAGAACCAAGGGGTGATCAGGAGATGGAGACTCTACAAAGAAAGAT 2916
QY 968 AspTyrGlnLysPheSerGlnLysLeuGlnGlyLysLysGlnProLysGlnProLysAla 987
Db 2917 GACTATAGGCTTTTCAGGTTCTCAGAAATTTGAAAGAAAGTTCAAACCAAAATTCCTGCG 2976
QY 988 SerSerMetProGlnGlnLysGlyLeuProThrSerSerLeuThrGlnValAlaHisSer 1007
Db 2977 TCTTTCATGCGCAACGAAATAGGCAATTCCTCAACAGTCTTTAACTCAAGTTGTTCAATTC 3036
QY 1008 AlaGlnValArgPheLeuValSerProValProGlnSerArgLeuArgGlnSerLysVal 1027
Db 3037 GCGGAGAGGGGTTTATAGTGAATCTCTGTCGCCAGAAAGCCGATTAAGAGAAATCAAAAGT 3096
QY 1028 PheProSerGlnLysThrAspThrValAlaAlaSerThrAlaGlnSerProGlnMetAsn 1047

```

```

Db 3097 TTCCCAAGAAATACAGATACATGCTGCTCTTACAGCTCAGAGCCCTGAAATGAAC 3156
QY 1048 LeuSerHisSerAlaSerSerLeuSerLeuGlnGlnAlaPheSerGlnLeuArgArgAla 1067
Db 3157 TTGTCTCACTCGATCATCTCTTATGCTTACACAGGCTTTTCTGAACTTATGAGCTGCTC 3216
QY 1068 GlnMetThrGlnGlyProAsnThrAlaProProAsnPheSerHisThrGlyProThrPhe 1087
Db 3217 CAATGACAGAGAGGCCCAAAVACAGACCTCCAAACTTAGTCATACAGAGCAACATTT 3276
QY 1088 ProValValProProPheLeuSerSerLysAlaGlyValProThrThrAlaAlaThr 1107
Db 3277 CCAATGATCCTCTTCTTATAGTACATGCTGTGAGTCCCAACACAGCAGCAGCAGCA 3336
QY 1108 AlaProValProAlaThrSerSerProProAsnAspLysSerThrSerValLysGlnSer 1127
Db 3337 GCACAGTCCCTGCAAGACAGAGAGCCCTCTTAATGACATTTCCACATCAGTAAATTCAGT 3396
QY 1128 GlnValThrValProThrGlnGlnGlyLysAlaGlyValAlaThrSerThrGlyValVal 1147
Db 3397 GAGTTACAGTGCCTCAGAGAGAGGAGATGCTGGAGTTGCCACAGCAGAGTGTGTA 3456
QY 1148 ThrSerGlyLysLeuProLysProLysProValSerGlnSerProValSerSerValVal 1167
Db 3457 ACTTCAGGTGCTTCCCATACACACTGTGTGTAAATCAACAGTACTTTCAGAGCGTAGT 3516
QY 1168 SerSerLysThrLysProAlaValValSerLysSerThrThrSerProSerLeuGlnVal 1187
Db 3517 TCAAGTATCAAAATACCTGAGGTTGTCTCAATATCTACTAATCCCGTCACTTCAGTGC 3576
QY 1188 ProThrSerThrSerGlnLysValValSerSerThrAlaLeuLysProSerValThrVal 1207
Db 3577 CCCACATCCACATGTGAGATCGTTGTTCTAGTACAGACAGTATCTTCACTGTAACAGT 3636
QY 1208 SerAlaThrSerAlaSerAlaGlyLysSerThrAlaThrProGlnProLysProProAla 1227
Db 3637 TCAGCACTTCAGCTCTGCGAGGGGAGATGATGCTACCCCAAGGTCTTAAGCCCTCAGCT 3696
QY 1228 ValValSerGlnGlnAlaAlaGlySerThrThrValGlyAlaThrLeuThrSerValSer 1247
Db 3697 GTAGTATCTCAGCAGCAGGAGGAGCAGCAGTACGTGGAGGACATTAATCATCAGTTCT 3756
QY 1248 ThrThrThrSerPheProSerThrAlaSerGlnLeuSerLysGlnLeuSerSerThr 1267
Db 3757 ACCACCACTTCATTTCCACACACAGCTTCAAGCTGTCCATTCAGCTTAGCAGCAGTACT 3816
QY 1268 SerThrProThrLeuAlaGlnThrValValValSerAlaHisSerLeuAspLysThrSer 1287
Db 3817 TCTACTCTCTACTTATGCTGAACCGGTGTAGTGGCGCACATCACTAGATMAAGACATCT 3876
QY 1288 HisSerSerThrThrGlyLeuAlaPheSerLeuSerAlaProSerSerSerSerPro 1307
Db 3877 CATAGCAGTACAACTGATGCTTCTCTCTCTGCAACATCTTCTCTCTCTCTCTCTCT 3936
QY 1308 GlyAlaGlyValSerSerTyrLysSerGlnProGlyGlyLeuHisProLeuValLysPro 1327
Db 3937 GGAGCAGAGATGCTGATTAATATTTCTCACCTGTGGGGCTGTGATCTTTGGTCAATCCA 3996
QY 1328 SerValLysLeuSerThrProLysLeuProGlnAlaAlaGlyProThrSerThrProLeu 1347
Db 3997 TCAGTATAGCTTCTACTCTCTATTTCTTCCCAAGCAGCAGAGACTTACTTACACCTTTA 4056
QY 1348 LeuProGlnValProSerLysLeuProLeuValGlnProValAlaAsnValProAlaVal 1367
Db 4057 TTACCCCAAGTACTTATGATCCACCTTGTGACAGCTGTGTCATGCTGTGCTGTGTA 4116
QY 1368 GlnGlnThrLeuLysHisSerGlnProGlnProAlaLeuLeuProAsnGlnProHisThr 1387
Db 4117 CAGCAGACACTAATTCATGTCAGCTCAACAGCTTGTCTTCCCAACAGCAGCAGCTACT 4176
QY 1388 HisCysProGlnValAspSerAspThrGlnProLysValProGlyLysAspAlaLys 1407

```

Db 4177 CATGTCCTGAGTAGATTCGTGATACCAACCCAAAGCTCTGGAAATTGATGACATAAG 4236
Qy 1408 ThrLeuGluGluLysLeuArgSerLeuPheSerGluHisSerSerSerGlyValGlnHis 1427
Db 4237 ACTCTGAAGAAAGAGCGCGCTCTGTCTGATGTAAGAACAGAGCTCATCTGAGCTCAGCAT 4296
Qy 1428 AlaSerValSerLeuGluThrSerLeuValIleGluSerThrValThrProGlyIlePro 1447
Db 4297 GCCTCTCTCACTCGAGAGACTCACTGATCATAGAGAGCACTGTGCACACAGGATCCCA 4356
Qy 1448 ThrThrAlaValAlaProSerLysLeuLeuThrSerThrThrSerThrCyLeuProPro 1467
Db 4357 ACTACTCTGTGGACCAAGCAAACTCTGACTTCAACCAAGTACTTCTTACCA 4416
Qy 1468 ThrAsnLeuProLeuGlyThrValAlaLeuProValThrProValAlaThrProGlyGln 1487
Db 4417 ACCAATTACCAACAGAAACAGTTGCTTGCAGATTACACCAAGTGCACACCTGGGCA 4476
Qy 1488 ValSerThrProValSerThrThrThrSerGlyValLysProGlyThrAlaProSerLys 1507
Db 4477 GTTTCACCCAGTCAGCACTACTACATCAGAGTGAACCTGSAACCTGCTCCCTCCAG 4536
Qy 1508 ProProLeuThrLysAlaProValLeuProValGlyThrGluLeuProAlaGlyThrLeu 1527
Db 4537 CCACCTCTAAGCTCCGCTCGCTGCTGAGTGGGTACTGAACTTCCAGCAGGTA 4596
Qy 1528 ProSerGluGluLeuProProPheProGlyProSerLeuThrGlnSerGlnGlnProLeu 1547
Db 4597 CCCAGCGCAGCTGCCACCTTTCAGAGACCTTCTAACCCAGTCCCAACCACTCTTA 4656
Qy 1548 GluAspLeuAspAlaGlnLeuArgArgThrLeuSerProGlu**IleThrValThrSer 1567
Db 4657 GAGGATCTTGAATGCTCAATTGAGAAAGAACCTTAGTCCAGAGATKATCAGCTACTTCT 4716
Qy 1568 AlaValGlyProValSerMetAlaAlaProThrAlaIleThrGlnAlaGlyThrGlnPro 1587
Db 4717 GCGGTGGTCTCTGTCATGCGCGCTCCCAACCAATCAGAAAGCAGAACAGCT 4776
Qy 1588 GlnLysGlyValSerGlnValLysGlnGlyProValLeuAlaThrSerSerGlyValGly 1607
Db 4777 CAGAGGGGTGTTTTCAGAGTAAAGAAAGCCCTGTCTAGCAACTGATTCAGAGAGTGT 4836
Qy 1608 ValPheLysMetGlyArgPheGlnValSerValAlaAlaAspGlyAlaGlnLysGlnGly 1627
Db 4837 GTTTTAAAGATGAGACGATTTCAAGTTCTGTGTGACGACGAGGCCCAAAAGAGGCT 4896
Qy 1628 LysAsnLysSerGluAspAlaLysSerValHisPheGluSerSerThrSerGluSerSer 1647
Db 4897 AAAAATAAGTCAGAAAGATGCAAAAGTCTGTTCATTTGAATCCAGCACTCAGAGCTCA 4956
Qy 1648 ValLeuSerSerSerSerProGluSerThrLeuValLysProGluProAsnGlyIleThr 1667
Db 4957 GTGCTATAGAGTAGTAGTCCAGAGAGTACTTGTGGAAACCAAGACCCAAATGGCAATACC 5016
Qy 1668 IleProGlyLysSerSerAspValProGluSerAlaHisLysThrThrAlaSerGluAla 1687
Db 5017 ATCCCTGGTATCTCTTCAGATGTGCCAGAGAGTGCACAAACTACTGCTCAGAGCA 5076
Qy 1688 LysSerAspThrGlyGlnProThrLysValGlyArgPheGlnValThrThrAlaAsn 1707
Db 5077 AAGTCAGACACTGGGACGCTCCACAGGTGACGTTTTCAGGTGACAACTCAGCAACAAAC 5136
Qy 1708 LysValGlyArgPheSerValSerLysThrGluAspLysIleThrAspThrLysLysGln 1727
Db 5137 AAAGTGGGTGCTTCTCTGTATCAAAAACCTGAGGACAGATCAGCAACAAAGAAAGAA 5196
Qy 1728 GlyProValAlaSerProPheMetAspLeuGlnAlaValLeuProAlaValIle 1747
Db 5197 GAGCAGATGGCATCTCTCTCTTTATGATTTGAAACAGCTGTCTCTCTGCTGATTA 5256
Qy 1748 ProLysLysGluLysProGluLeuSerGluProSerHisLeuAsnGlyProSerSerAsp 1767
Db 5257 CCAAGAAAGAGAGAGCTGAACTGTACAGGCTTCACTTAATGGGCGCTCTCTGTAC 5316

Qy 1768 ProGluAlaAlaPheLeuSerArgAspValAspAspGlySerGlySerProHisSerPro 1787
Db 5317 CCGAGAGCGCGCTTTTAAATAGAGATGTGAGATGATGATTCGCGTATGCACACTCCGCC 5376
Qy 1788 HisGlnLeuSerSerLysSerLeuProSerGlnAsnLeuSerGlnSerLeuSerAsnSer 1807
Db 5377 CATCAGCTGAGCTCAAGAGCCCTTCTAGCCAAATCTTAAGTCAAAAGCTTATGATTA 5436
Qy 1808 PheAsnSerSerThrMetSerSerAspAsnGluSerAspIleGluAspGluAspLeuLys 1827
Db 5437 TTAACTCTCTTAATGATGATGACGACATGATGATGATTCGAAATGAAAGCTTAAG 5496
Qy 1828 LeuGluLeuArgArgLeuArgAspLysHisLeuLysGluIleGlnAspLeuGlnSerArg 1847
Db 5497 TTAGAGCTGCGAGCACTACGAGATTAACATCTAAAGAGATTCAGAGCTGAGAGTCCG 5556
Qy 1848 GlnLysHisGluIleGluSerLeuThrLysLeuGlyLysValProProAlaValIle 1867
Db 5557 CAGAGAGATGAATAATTGAAATCTTGTATACCAAACTGGGCAAGGTGCCCTGCTGTATT 5616
Qy 1868 IleProProAlaAlaProLeuSerGlyArgArgArgProThrLysSerLysGlySer 1887
Db 5617 ATTCCCTCAGCTGCTCCCTTTCAAGAGAAAGACGACACCTTAAGCAAGGAC 5676
Qy 1888 LysSerSerArgSerSerSerLeuGlyAsnLysSerProGluLeuSerGlyAsnLeuSer 1907
Db 5677 AAATCTAGTGAAGCAGATTCCTTGGGGAATAAAGCCCCAGCTTCAAGTAACTGTCT 5736
Qy 1908 GlyGlnSerAlaAlaSerValLeuHisProGlnGlnThrLeuHisProProGlyAsnIle 1927
Db 5737 GGTACAGATCAGCTTCACTTGGCACCCAGAGAGCCCTCCACCTCTCGGAAATC 5796
Qy 1928 ProGluSerGlyGlnAsnGlnLeuLeuGlnProLeuLysProSerProSerSerAspAsn 1947
Db 5797 CCAGAGTCGGGAGAAATCAGCTGTATACAGCCCTTAAGCATCTCTCCAGAGTACAC 5856
Qy 1948 LeuThrSerAlaPheThrSerSerAspGlyAlaIleSerValProSerLeuSerAlaProGly 1967
Db 5857 CTCTATTCAGCTTCAACAGTATGTCATTCATTCAGTAACTTCTGCTCCAGT 5916
Qy 1968 GlnGlyThrSerSerThrAsnThrValGlyAlaThrValAsnSerGlnAlaAlaGlnAla 1987
Db 5917 CAAGAAACAGACACAAACACTGTGGGCAACAGTGAACACCAAGCGCCCAAGCT 5976
Qy 1988 GlnProProAlaMetThrSerSerArgLysGlyThrPheThrAspAspLeuHisLysLeu 2007
Db 5977 CAGCTCTGTCATGAGTCCAGAGAGAGGAGCATTCACAGATGACTTGCACAAATG 6036
Qy 2008 ValAspAsnTrpAlaArgAspAlaMetAsnLeuSerGlyArgArgGlySerLysGlyHis 2027
Db 6037 GTAGACAAATGGGCGCCAGAGTGCATGATCTTCAGGCAAGAGAGAAAGGAGC 6096
Qy 2028 MetAsnTrpGlnGlyProGlyMetAlaArgLysPheSerAlaProGlyGlnLeuCyIle 2047
Db 6097 ATGAATTATGAGGCGCTGGAATGCAAGAAATTCCTGCACTGGGCAACTGTGATC 6156
Qy 2048 SerMetThrSerAsnLeuGlyGlySerAlaProIleSerAlaAlaSerAlaThrSerLeu 2067
Db 6157 TCCATGACCTCGAACCTGAGTGGCTTCCCTCCATCTTGCAGCATGCACTTCTCTTA 6216
Qy 2068 GlyHisPheThrLysSerMetCysProProGlnGlnGlyPheProAlaThrProPhe 2087
Db 6217 GGTCACTTCAACCAATGATATGAGCCCCCAAGCAGTATGCTTCCAGTAACTCCATTT 6276
Qy 2088 GlyAlaGlnTrpSerGlyThrGlyGlyProAlaProGlnProLeuGlyGlnPheGlnPro 2107
Db 6277 GCGCTCATGAGAGTGGAGCGGTGGCCAGACCAAGCACTTGGCCAGTTCCAACT 6336
Qy 2108 ValGlyThrAlaSerLeuGlnAsnPheAsnIleSerAsnLeuGlnLysSerIleSerAsn 2127
Db 6337 GTGGAACTGCTCTTGCAGATTTCAACATCAGCAATTTTGCAGAAATTCATGCAAC 6396

QY 2128 ProProGlySerAsnLeuArgThrThr 2136
 DB 6397 CCCCAGGCTCCAACTGCGGACCACT 6423
 RESULT 15
 ACH03788
 ID ACH03788 standard; DNA; 6885 BP.
 AC ACH03788;
 XX ACH03788;
 AC ACH03788;
 DT 26-SEP-2003 (first entry)
 XX DNA encoding novel human protein kinase #25.
 DE Human; kinase; ds; gene; cosmetic application; nutraceutical application.
 XX Homo sapiens.
 OS Homo sapiens.
 XX US6541252-B1.
 XX 01-APR-2003.
 XX 14-MAY-2001; 2001US-00854856.
 XX 19-MAY-2000; 2000US-0206015P.
 XX (LEXI-) LEXICON GENETICS INC.
 XX Walke DW, Hilbun E, Donoho G, Turner CA;
 XX WPI: 2003-575927/54.
 XX P-PSDB; ABO44405.
 PT New nucleic acid encoding novel human proteins, useful in cosmetic and
 PT nutraceutical applications.
 XX PS Disclosure; Page; 11pp; English.
 CC The invention relates to a new isolated nucleic acid encoding a novel
 CC human protein kinase. The nucleic acid is useful in cosmetic and
 CC nutraceutical applications. The present sequence represents DNA encoding
 CC a novel human protein kinase. Note: The sequence data for this patent did
 CC not from part of the pinned specification but was obtained in electronic
 CC format directly from USPTO at
 CC seqdata.uspto.gov/sequence.html?docid=6541252B1
 XX SQ Sequence 6885 BP; 1881 A; 1873 C; 1543 G; 1585 T; 0 U; 3 Other;
 Alignment Scores:
 Pred. No.: 0 Length: 6885
 Score: 10211.00 Matches: 2048
 Percent Similarity: 88.20% Conservative: 0
 Best Local Similarity: 88.20% Mismatches: 0
 Query Match: 94.44% Indels: 274
 DB: 8 Gaps: 2
 US-10-010-720-14 (1-2136) x ACH03788 (1-6885)
 QY 61 MetAspLysAspSerArgGlyAlaAlaAlaThrThrThrThrThrGlnHisArgPhePhe 80
 DB 1 ATGAGCAAGACAGACCGCTGGGGCGGCGGACCACTACCACTGAGCAACCGCTTCTTC 60
 QY 81 ArgArgSerValIleCysAspSerAsnAlaThrAlaLeuGluLeuProGlyLeuProLeu 100
 DB 61 CGCCGAGGCTCATCTGCGACTCCAAATGCCACTGAGAGCTTCCGCGCTTCTCTT 120
 QY 101 SerLeuProGlnProSerIleProAlaAlaValProGlnSerAlaProProGlnProHis 120
 DB 121 TCCCTGCCCAAGCCAGCATCCCGCGGCTGTCGCGAGAGTCTCCACCGGAGCCCAAC 180
 QY 121 ArgGluGluThrValThrAlaThrAlaThrSerGlnValAlaGlnGlnProProAlaAla 140
 DB 181 CGGGAAGACACCTGACCGCCACCGCACTTCCAGAGTAGCCCAAGAGCTCCAGCCGCT 240

QY 141 AlaAlaProGlyGluGlnAlaValAlaGlyProAlaProSerThrValProSerSerThr 160
 DB 241 GCGGCCCTGGGGAGACAGCCGCTCGCGGCGCTCCCTCGACTGTGCTCCAGCAGTACC 300
 QY 161 SerLysAspArgProValSerGlnProSerLeuValGlySerLysGluGluProProPro 180
 DB 301 AGCAAAAGACCGCCAGGTCTCCAGCTTAGCCTTGTGGGAGGCAAGAGAGCGCGCCG 360
 QY 181 AlaArgSerGlySerGlyGlySerAlaLysGlnProGlnGluGluArgSerGlnGln 200
 DB 361 GCGAAGAGTGGCAGCGCGCGCGGACGCCCAAGAGCCACAGAGAGAGAGAGAGAGAG 420
 QY 201 GlnAspAspIleGluGluLeuGluThrLysAlaValAlaGlyMetSerAsnAspGlyArgPhe 220
 DB 421 CAGATGATATCGAAGACCTGAGACCAAGCCGCGGAGATGCTTCAAGATGCGCGCTT 480
 QY 221 LeuLysPheAspIleGluIleGlyArgGlySerPheLysThrValTyrLysGlyLeuAsp 240
 DB 481 CTCAAGTTGACATCGAATGCGCAGAGCTCTTTAAGAGGTCTACAAAGGTCTGAGC 540
 QY 241 ThrGluThrThrValGluValAlaTyrCysGluLeuGlnAspArgLysLeuThrLysSer 260
 DB 541 ACTGAAACCAACCGTGAAGTCCGCTGCTGTAATGCAAGATCAAAATTAACAAAGCT 600
 QY 261 GluArgGlnArgPheLysGluGluAlaGluMetLeuLysGlyLeuGlnHisProAsnIle 280
 DB 601 GAGAGGACAGATTTAAAGAAAGAGCTGAATGTTAAAGGTCTTCAACATCCCAATATT 660
 QY 281 ValArgPheTyrAspSerTyrPylSerThrValLysGlyLysCysIleValIleVal 300
 DB 661 GTTAGATTATATGATTCCTGGGAATCCACAGTAAAGAAAGAGTGTATTGTTG 720
 QY 301 ThrGluLeuMetThrSerGlyThrLeuLysThrTyrLeuLysArgPheLysValMetLys 320
 DB 721 ACTGAACCTTAGACGTGGAACACCTTAAACGATCTGAAGAGTTTAAAGTATGATGAG 780
 QY 321 IleLysValLeuArgSerTyrPylSerThrValLysGlyLysGlyLeuGlnHisPheHisThr 340
 DB 781 ATCAAACTTCAAGAACTGCTGCTGCAATCTTAAAGTCTTCAAGTCTTCAATCT 840
 QY 341 ArgThrProProIleIleHisArgAspLeuLysCysAspAsnIlePheIleThrGlyPro 360
 DB 841 CGAATCCACCTTCAATCAATCCGGAATCTTAATGTACAAATCTTATCACCAGGCT 900
 QY 361 ThrGlySerValLysIleGlyAspLeuGlyLeuAlaThrLeuLysArgAlaSerPheAla 380
 DB 901 ACTGGCTCAGTCAAGATTGAGACCTCGTCTGGCAACCTGAAGCGGCTTCTTTGGCC 960
 QY 381 LysSerValIleGlyThrProGluPheMetAlaProGluMetTyrGluGluLysTyrAsp 400
 DB 961 AAGAGTGTATAGTACCCCAAGATTCATAGCCCTGAGATGATGAGGAAATATAGAT 1020
 QY 401 GluSerValAspValTyrAlaPheGlyMetCysMetLeuGlnMetAlaThrSerGluTyr 420
 DB 1021 GAATCGTTGACGTTTATGCTTTTGGGATGTGATGCTTATGATGCTTATGATAT 1080
 QY 421 ProTyrSerGluCysGlnAsnAlaAlaGlnIleTyrArgArgValThrSerGlyValLys 440
 DB 1081 CCTTACTCGGAGTCCAAATGCTGCACAGATCTACCGTGCCTGACCACTGGGGGTGAAG 1140
 QY 441 ProLysSerPheAspLysValAlaIleProGluValLysGluIleGluGlyCysIle 460
 DB 1141 CAGGCCAGTGTGACAAAGTAGCAATCTCGAAGTAGAAGAAATTTTGAAGATGCATCA 1200
 QY 461 ArgGlnAsnLysAspGluArgTyrSerIleLysAspLeuLeuAsnHisAlaPhePheGln 480
 DB 1201 CGACAAAGACAAAGATGAAAGATATTCATCAAGACCTTTTGAACCATGCTTCTTCCAA 1260
 QY 481 GluGluThrGlyValArgValGluLeuAlaGluGluAspGlyGluLysIleAlaIle 500
 DB 1261 GAGGAACAGAGATGAGGCTAGATTAAGCAAGAGATGATGAGAAATAATAGCATAT 1320

QY 501 LysLeuTrpLeuArg11egluAsp11elysLeuLysGlyLysTyLysAspAsnGlu 520
 |||||
 Db 1321 AAATTAATGCTAGTATTAAGATATTAAGAAATTAAGGAAATTAACAAGATATTAAGA 1380
 QY 521 Ala11leGluPheSerPheAspLeuGluArgAspValProGluAspValAlaGlnGluMet 540
 |||||
 Db 1381 GCTATGAGTCTTCTTTGATTTAAGAGAGATGTCCTCAAGAAATGTTGACAAAGAAATG 1440
 QY 541 ValGluSerGlyTyTrValGlyGluLysAspHisLysThrMetAlaLysAla11elysAsp 560
 |||||
 Db 1441 GTAGAGTCTGGTATGCTGTGAAGGTGATCAAGACCAATGGCTAAAGCTATCAAGAC 1500
 QY 561 ArgValSerLeuLeuLysArgLysArgGluGlnArgGlnLeuValArgGluGlnGlu 580
 |||||
 Db 1501 AGAGTATCATTAATTAAGAGAAACGAGACAGCGGAGTGTGTAGCGAGAGAGAA 1560
 QY 581 LysLysLysGlnGluGluSerSerLeuLysGlnGlnValGluGluSerSerAlaSerGln 600
 |||||
 Db 1561 AAAAAAAGCAGAGAGAGACAGTCTCAACAGCAGAGTAGAACATCCAGTGTCCAG 1620
 QY 601 ThrGlyLysGlnLeuProSerAlaSerThrGlyLysProThrAlaSerThrThrSer 620
 |||||
 Db 1621 ACGAGATCAAGACGCTCCCTCTGTAGACACGGCATACCTACTGCTTACCACTTCA 1680
 QY 621 AlaSerValSerThrGlnValGluProGluGluProGluAlaAspGlnHisGlnGlnLeu 640
 |||||
 Db 1681 GCTTCACTTTCTACAGATGAAGTGAACCTGAAGAACTGAGGAGATCAACATCAACATCA 1740
 QY 641 GlnTyGlnGlnProSer11SerValLeuSerAspGlyTyTrValAspSerGlyGlnGly 660
 |||||
 Db 1741 CAGTACAGCAACCCAGTATATCTGTATCTGATGAGAGCGGTGACAGTGTGACAGGGA 1800
 QY 661 SerSerValPheThrGluSerArgValSerSerGlnGlnThrValSerTyGlySerGln 680
 |||||
 Db 1801 TCCCTCTGTCTTCAAGATCTGAGTAGAGAGCCAGCAAGACATTTATATATGTTCCCA 1860
 QY 681 HisGlnGlnAlaHisSerThrGlyTyTrValProGlyHisLysProSerThrValGlnAla 700
 |||||
 Db 1861 CATAAACAGGACATTTCTACAGGACAGTCTCCAGGCAATATACCTTCTACTGTCCAGCA 1920
 QY 701 GlnSerGlnProHisGlyValTyTrProProSerSerValAlaGlnGlyGlnSerGlnGly 720
 |||||
 Db 1921 CAGTCTCAGCCCATGGGGTATATCCACCTCAAGTGTG----- 1959
 QY 721 GlnProSerSerSerSerLeuThrGlyValSerSerSerGlnProLysGlnHisProGln 740
 |||||
 Db 1959 ----- 1959
 QY 741 GlnGlnGlnGlyLysGlnGlnThrAlaProProGlnGlnThrValGlnTySerLeuSer 760
 |||||
 Db 1960 ---CAGCAGGGAATACAGACACAGACCCCTCTCTCAACAGACAGTGCATATTCATCTTCA 2016
 QY 761 GlnThrSerThrSerSerGlyAlaThrThrAlaGlnProValSerGlnProGlnAlaPro 780
 |||||
 Db 2017 CAGACATCAACCTCCAGTAGGAGCCACTACTGACACAGCAGTAGAGTCAAGCTCAAGTCCA 2076
 QY 781 GlnValLeuProGlnValSerAlaGlyLysGln----- 791
 |||||
 Db 2077 CAAGCTTGCCTCAAGTATCAAGCTGGAAGAAAGAGCTTCCAGTTCCACAGCAATCAACT 2136
 QY 791 ----- 791
 Db 2137 ATCCAGGCGAACTCAATCCAGTCCGACACAAACCTCGGTTGTTCCAGTCCACT 2196
 QY 791 ----- 791
 Db 2197 GATGCTCATTTCTTCCAGTGGAGACAGCGCTCTCTACTCCCTTGCCTCAGTACCT 2256
 QY 791 ----- 791
 Db 2257 GTCTCTCAGATTTCCATATCAACTCTCTATGTGTCTACGGCTCAGACAGGTTTCTCATCC 2316
 QY 791 ----- 791

Db 2317 CTTCCTCATCAATGACAGCTGGCATTTACTCAGCCTTGTCTACGTTGCTTCACTGCT 2376
 QY 791 ----- 791
 Db 2377 ACAACAGCTGGCATTCGGGGGGTATCACTGTGTCTTAAGTACGTTCCAACTTCTCTG 2436
 QY 791 ----- 791
 Db 2437 CAGCTGTAGCTAGCTGCCAAGTCAAGTTCACCCAGCTCTCAACACAGCAGATTACAG 2496
 QY 791 ----- 791
 Db 2497 TCCATGGAATACCACTTAACCTTGGACCAAGCTGTGAAGTTCACATTTCTCTGAGAT 2556
 QY 791 ----- 791
 Db 2557 GTTCTGTACAGGGCTTCCCACTCGACTGCCACACAGTACCCAGAGATTCAATATT 2616
 QY 791 ----- 791
 Db 2617 GCTCCCTCTTCAACGTGAGTCTGTGTTGATCATCTTCAAGTCTTACGCCCTCCATG 2676
 QY 791 ----- 791
 Db 2677 CCGACAGAGTACTGCTACACTGGGTACTTCCCAAGTGTGACCTTATGTGAA 2736
 QY 791 ----- 791
 Db 2737 TCAATTTTATTTAGTCTTATGAGGTGTGTAGAGACAGGTCAAGTGTCCAGCCAGCA 2796
 QY 792 -----SerThrGln 794
 Db 2797 GGGAGTTAGACAAAGCCCACTACATCTCCACAGACAGATTTGAGAGTACTCAG 2856
 QY 795 GlyValSerGlnValAlaProAlaGluProValAlaValAlaGlnProGlnAlaThrGln 814
 |||||
 Db 2857 GGAATCTCTCAGGTGCTCTCTGACAGAGCAGTTGACATAGACAGCCCAAGTACCCAG 2916
 QY 815 ProThrThrLeuAlaSerSerValAspSerAlaHisSerAspValAlaSerGlyMetSer 834
 |||||
 Db 2917 CCGACCACTTGTGGTCTCTCTGTAGACAGTGCATTCAGATGTGCTTCCAGTATGAGT 2976
 QY 835 AspGlyAsnGluAsnValProSerSerSerGlyArgHisGlyGlyArgThrThrLysArg 854
 |||||
 Db 2977 GATGCAATGAAAGTCCCATCTTCCAGTGAAGGATGAAGAGAACTTCACAAAACGG 3036
 QY 855 HisTyArgLysSerValArgSerArgSerArgHisGlyLysThrSerArgProLysLeu 874
 |||||
 Db 3037 CATTAACGAAATCTGTAGAGAGTGCCTTCGACATGAAAAAACTTCAAGCCCAAAATTA 3096
 QY 875 ArgGluLeuAsnValSerAsnLysGlyAspArgValAlaGluCysGlnLeuGluThrHis 894
 |||||
 Db 3097 AGAATTTTGAATGTTTCAATATTAAGAGACCGAGTAGTAGATTCATTTAGACATCAT 3156
 QY 895 AsnArgLysMetValThrPheLysPheAspLeuAspGlyAspAsnProGluGluLeuAla 914
 |||||
 Db 3157 AATAGGAATAGTTTCAATATTAAGAGACCGAGTAGTAGATTCATTTAGACATCAT 3216
 QY 915 ThrLysMetValAsnAspPheLysLeuAla11egluArgGluSerPheValAspGln 934
 |||||
 Db 3217 ACAATTTATGTGAACAATGACTTTATCTTACCAATAGAGAGAGTGGTTGTGATCAA 3276
 QY 935 ValArgGluLeu11egluLysAlaAspGluMetLeuSerGluAspValSerValGluPro 954
 |||||
 Db 3277 GTGGCAGAAATTTATTAAGAAAGCTGATGAATATCTCTCAGTACGATGTCAAGTGAACA 3336
 QY 955 GluGlyAspGlnGlyLeuGluSerLeuGlnGlyLysAspAspTyGlyPheSerGlySer 974
 |||||
 Db 3337 GAGGGTATTCAGGGATTTGAAGAGTCTCAAGAGAAAGATGACTTATGCTTTTCAAGTCT 3396
 QY 975 GlnLysLeuGluGluPheLysGlnProLysSerMetProGlnGlnLeu 994
 |||||

3397 CAGAAATTGAGAGAGAGTTCACCAACCAATTCCTGCGTTCCTCATGCCACAGCAATA 3456
QY 995 GYLIEProThrSerSerLeuThrGlnValHisSerAlaGlyArgArgPheIleVal 1014
Db 3457 GGCATTCCTACCAAGTTCCTTAACGATCAAGTGTTCATTCGCGGAAAGCGGTTTAATG 3516
QY 1015 SerProValProGluSerArgLeuArgGluSerValPheProSerGluIleThrAsp 1034
Db 3517 AGTCCGTGCGCAGAAAGCCGATTCAGAGATCAAAAGTTTCCCAAGTAATACAT 3576
QY 1035 ThrValAlaAlaSerThrAlaGlnSerProGlyMetAsnLeuSerHisSerAlaSerSer 1054
Db 3577 ACAGTGTGCTCTACAGCTCAGAGCCCTGGAAGAACTGTCTCACCTGCAATCC 3636
QY 1055 LeuSerLeuGlnGlnAlaPheSerGluLeuArgArgAlaGlnMetThrGluGlyProAsn 1074
Db 3637 CTTAGTCTACACAGCCCTTTTCTGAACTTAGACCTGCCCAAAAGACAGAAAGACCAAY 3696
QY 1075 ThrAlaProProAsnPheSerHisThrGlyProThrPheProValValProPheLeu 1094
Db 3697 ACAGCACCCTCCAACTTAGTCATACAGACCAACATTTCCAGTAGTACCTCTTCTTA 3756
QY 1095 SerSerIleAlaGlyValProThrThrAlaAlaAlaThrAlaProValProAlaThrSer 1114
Db 3757 AGTACGATGCTGAGAGTCCACACACAGACAGCCACAGCACAGCCTCCGCAACAGC 3816
QY 1115 SerProProAsnAspIleSerThrSerValIleGlnSerGluValThrValProThrGlu 1134
Db 3817 AGCCCTCCTAATGACATTTCCACATCAGTAATTCAGTCTGAGTTACGTCCCACTGAA 3876
QY 1135 GlnGlyIleAlaGlyValAlaThrSerThrGlyValValThrSerGlyLeuProIle 1154
Db 3877 GAGGGGATGCTGAGTGCACACACAGCAGGTGTGTAACTTCAGAGGTGCTCCCA 3936
QY 1155 ProProValSerGluSerProValLeuSerSerValValSerSerIleThrIleProAla 1174
Db 3937 CCACCTGTGTGATCAACAGTACTTCCAGCTGTAAGTTCMAATATCACAAATCTCGCA 3996
QY 1175 ValValSerIleSerThrThrSerProSerLeuGlnValProThrSerThrSerGluIle 1194
Db 3997 GTTGTCTCAATATCTACTACATCCCGTCACTTAAGTCCCAATCCACATCTTGAAATC 4056
QY 1195 ValValSerSerThrAlaLeuTyProSerValThrValSerAlaThrSerAlaSerAla 1214
Db 4057 GTTGTCTCTAGTACAGACTGTATCTTCAGTACAGATTCCAGCACTTCAGCCTCGCA 4116
QY 1215 GlyGlySerThrAlaThrProGlyProValProProAlaValValSerGlnAlaAla 1234
Db 4117 GGGGGCAGTACTGCTACCCAGGCTTAAAGCTCCACCTGATATCTCAGCGAGCA 4176
QY 1235 GlySerThrThrValGlyAlaThrLeuThrSerValSerThrThrThrSerPheProSer 1254
Db 4177 GGCAGCACTACTGTGGAGCCACTTAACATCAGTTTCTACACCACTTCAATCCCAAGC 4236
QY 1255 ThrAlaSerGlnLeuSerIleGlnLeuSerSerSerThrSerThrProThrLeuAlaGlu 1274
Db 4237 ACAGCTTCACAGCTGTCCATTCAGCTTACAGCAGTACTTCTACTCTTAAGTGAA 4296
QY 1275 ThrValValValSerAlaHisSerLeuAspLeuThrSerHisSerSerThrThrGlyLeu 1294
Db 4297 ACCGTGTAGTTAGCGCACACTCACTAGTAAACATCTATAGCAGTAACTGGAATTG 4356
QY 1295 AlaPheSerLeuSerAlaProSerSerSerSerSerProGlyAlaGlyValSerSerTy 1314
Db 4357 GCTTCTCCCTCTGTGACATCTTCTCTCTCTCTGAGACAGAGAGTCTTAGTAT 4416
QY 1315 IleSerGlnProGlyGlyLeuHisProLeuValIleProSerValIleAlaSerThrPro 1334
Db 4417 ATTTCACAGCTGTGGTGGTCACTTTGGTCAATCCATGATGATGATCTTACTCCGT 4476
QY 1335 IleLeuProGlnAlaAlaGlyProThrThrSerThrProLeuLeuProGlnValProSerIle 1354
Db 4477 ATTCTTCCCAAGCAGACGACTTCTTACACCTTTATATACCCCAAGTACTTACTATC 4536
QY 1355 ProProLeuValGlnProValAlaAsnValProAlaValGlnGlnThrLeuIleHisSer 1374
Db 4537 CCACCTTGTGTACAGCCTGTGGCCAAATGTGCTGTGACAGACACACTAATTCATGT 4596
QY 1375 GlnProGlnProAlaLeuLeuProAsnGlnProHisSerHisSerProGluValAspSer 1394
Db 4597 CAGCTTCACACAGCTTGTCTTCCCAACAGGCCCACTCATTTGTCTGGAAGTAAATCT 4656
QY 1395 AspThrGlnProValAlaProGlyIleAspAspIleTySerThrLeuGluGluLeuArg 1414
Db 4657 GATACACACCCCAAGCTCCTGAATGATGACATTAAGACCTTAGAAGAAAGCTGGG 4716
QY 1415 SerLeuPheSerGluHisSerSerSerGlyValaGlnHisAlaSerValSerLeuGluThr 1434
Db 4717 TCTGTGTCAGTGAACACAGCTCATCTGAGACTGAGCATGCTGCTGTCTCACTGGAGACC 4776
QY 1435 SerLeuValIleGlnSerThrValThrProGlyIleProThrThrAlaValAlaProSer 1454
Db 4777 TCACCTAGTCATAGAGACCTGTACACACAGGCAATCCCACTACTGCTGTGACCAAGC 4836
QY 1455 LysLeuLeuThrSerThrThrThrSerThrCysLeuProProThrAsnLeuProLeuGlyThr 1474
Db 4837 AAACCTCGACTTCTACCAACAAGTACTTGTACCAACAACATTTACACTAGAACCA 4896
QY 1475 ValAlaLeuProValThrProValValThrProGlyGlnValSerThrProValSerThr 1494
Db 4897 GTTGTCTTGCAGTTACACACAGTGTACACCTGTGGCAAGTTTCTACCCCACTGAGACT 4956
QY 1495 ThrThrSerGlyValLysProGlyThrAlaProSerLysProProLeuThrLysAlaPro 1514
Db 4957 ACTACATCAGAGAGGAAGCTGAGACTGCTCTCCCAAGCACCCTTAAAGCTCGC 5016
QY 1515 ValLeuProValGlyThrGluLeuProAlaGlyThrLeuProSerGlnGlnLeuProPro 1534
Db 5017 GTGTGTCAGTGGGTACTGAATCTCCAGCAGTACTTACCAAGCAGAGCTGCCACT 5076
QY 1535 PheProGlyProSerLeuThrGlnSerGlnProLeuGlnAspLeuAspAlaGlnLeu 1554
Db 5077 TTTCCAGACCTTCTCTAACCCAGTCCCAAGCACTCTTAAGAGATCTTGATGTCTCAATTG 5136
QY 1555 ArgArgThrLeuSerProGlu***IleThrValThrSerAlaValGlyProValSerMet 1574
Db 5137 AGAAGAACACTTAGTCCAGAGATKATCACGTGACTTCTGCGTTGCTCTGATCATG 5196
QY 1575 AlaAlaProThrAlaIleThrGlnAlaGlyThrGlnProGlnLysGlyValSerGlnVal 1594
Db 5197 GCGGCTCCAAACAGCAATCACAGAAAGCAGAAACAAGCTCAGAAAGGTGTCTCAAGTC 5256
QY 1595 LysGluGlyProValLeuAlaThrSerSerGlyAlaGlyValPheLysMetGlyArgPhe 1614
Db 5257 AAAAGAGGCCCTCTCTAGCACTAGTTCAGAGCTGTGTGTTTAAAGAGGAGCAATTT 5316
QY 1615 GlnValSerValAlaAlaAspGlyAlaGlnLysGlyLysAsnLysSerGluAspAla 1634
Db 5317 CAGTTTCTGTTCACACAGACGCTGCCAGAAAGAGGTAATAATAGTAAAGATGCA 5376
QY 1635 LysSerValHisPheGluSerThrThrSerGluSerSerValLeuSerSerSerPro 1654
Db 5377 AAGTGTTCATTTGAATCCAGCACTCAGAGTCTCAGGTGTCTTCAAGTATGATGCCA 5436
QY 1655 GluSerThrLeuValLysProGluProAlaGlyIleThrIleProGlyIleSerSerAsp 1674
Db 5437 GAGAGTACCTTGTGTGAACCCAGAGCCGAATGGATTAACATCCCTGTATCTCTTAGAT 5496
QY 1675 ValProGluSerAlaHisLysThrThrAlaSerGluAlaLysSerAspThrGlyGlnPro 1694
Db 5497 GTGCCAGAGAGTCCCAAAATTAATCTGCTCAGAGGCAAAAGTCAACACTGGGAGCT 5556
QY 1695 ThrLysValGlyArgPheGlnValThrThrThrAlaAsnLysValGlyArgPheSerVal 1714
Db 5557 ACCAAGGTGAGAGTTTTCAGGTGACAACTACAGCAAAAGAGTGGTGTCTCTGTGA 5616

QY 1715 SerLySThrGlnuAspLysIleThrAspThrIlySylGluGlyProValAlaSerProPro 1734
 Db 5617 TCAAAACAGTGAAGACCAAGATCACTGACACAAAGAAAGAGACAGGGGATCTCTCTCT 5676
 QY 1735 PheMetAspLeuGlnGlnAlaValLeuProAlaValIleProLysLysGlyLysProGlu 1754
 Db 5677 TTTATGGATTGGAAACAGCTGTTCTTCTGCTGATACCAAGAAAGAGAAAGGCTTGA 5736
 QY 1755 LeuSerGlnProSerIleuAsnGlnProSerSerAspProGlnAlaIlePheLeuSer 1774
 Db 5737 CTGTGAGAGCTTTCACATCTAAATGAGCGCTTCTTGACCCGAGGCGCTTTTAAAGT 5796
 QY 1775 ArgAspValAspAspGlySerGlySerProIleSerProIleGlnLeuSerSerLysSer 1794
 Db 5797 AGGAGATGTGATATGTTGTTCCGTAAGTCCACACTCCGCCCATGAGCTGAGCTCAAAGAGC 5856
 QY 1795 LeuProSerGlnAsnLeuSerGlnSerLeuSerAsnSerPheAsnSerSerLysMetSer 1814
 Db 5857 CTTCCTAGCCAGATCTAAAGCTAAAGCCCTTAAATTAATTCATTTAACTCTTACATGAGT 5916
 QY 1815 SerAspAsnGlnSerAspIleGlnuAspGlnuAspLeuLysLeuArgArgLeuArg 1834
 Db 5917 AGGACAAATAGTCAGATATCGAAGATGAAGACTTAAAGTTAAGCTGCGACGACTAGCA 5976
 QY 1835 AspLysIleLeuLysGlnIleGlnAspLeuGlnSerArgGlnLysIleGlnIleGlnSer 1854
 Db 5977 GATTAACATCTCAAAAGATTTCAAGACCTGACAGTCCAGAAAGCATGAATTGAATCT 6036
 QY 1855 LeuLysThrLysLeuGlnLysValProProAlaValIleIleProProAlaAlaProLeu 1874
 Db 6037 TTGATATCCAAACTGGGCAAGTGCCTCCCTGCTGTATTAATCCCCAGCTGCTCCCTT 6096
 QY 1875 SerGlyArgArgArgArgProThrLysSerLysGlySerLysSerSerArgSerSerSer 1894
 Db 6097 TCAGGAGAAAGACAGACCCACCTAAAGCAAAAGCAGCAAAATCTAGTCGAGAGCATTTCC 6156
 QY 1895 LeuGlyAsnLysSerProGlnLeuSerGlyAsnLeuSerGlyGlnSerAlaAlaSerVal 1914
 Db 6157 TTGGGGAATTAAGACCCCGACCTTTCAGGTAACTGTGCTGTCAGAGTGCAGCTTCACTC 6216
 QY 1915 LeuHisProGlnGlnThrLeuHisProProGlyAsnIleProGlnSerGlyGlnAsnGln 1934
 Db 6217 TTGCACCCCGACAGACCTTCACCTCTGCAACATCCCAAGTCCGGGAGCAATCAG 6276
 QY 1935 LeuLeuGlnProLeuLysProSerProSerSerAspAsnLeuLysSerAlaPheThrSer 1954
 Db 6277 CTGTTACAGCCCTTAAGCCATCTCCCTCAGTACAACTCTTAATCAGCTTCAACAGT 6336
 QY 1955 AspGlyAlaIleSerValProSerLeuSerAlaProGlyGlnGlyThrSerSerThrAsn 1974
 Db 6337 GATGCTGCATTTCAAGTACCAAGCTTCTGCTCAGGTCAAGAACAGCAGCAAAAC 6396
 QY 1975 ThrValGlyAlaThrValAsnSerGlnAlaIleGlnAlaGlnProProAlaMetThrSer 1994
 Db 6397 ACTGTTGGGCAACAGTGAACAGCCCAAGCCCAAGCTCAGCTCCGTCATGACGTCC 6456
 QY 1995 SerArgLysGlyThrPheThrAspAspLeuHisLysLeuValAspAsnThrAlaArgAsp 2014
 Db 6457 AGCAGAGAGGAGCACTTCAAGATGACTTGCACAACTGTGAGCAATTTGGGCGCGAGAT 6516
 QY 2015 AlaMetAsnLeuSerGlyArgArgLysSerLysGlyHisMetAsnLysGlyLysProGly 2034
 Db 6517 GCCATGATCTCTCAGGAGAGAGAGAGCAAAAGGCACTGAATTAAGAGGCGCTTGA 6576
 QY 2035 MetAlaArgLysPheSerAlaProGlyGlnLeuLysLysMetThrSerAsnLeuGly 2054
 Db 6577 ATGCAAGAGAGTCTCTGCAACTGGCAACTGTCATCTCATGACCTCCGAACCTGGGT 6636
 QY 2055 GlySerAlaProIleSerAlaAlaSerAlaThrSerLeuGlnHisPheThrLysSerMet 2074
 Db 6637 GGCTCTGCCCATCTCTGCAAGATCAGTACCTCTTAAGTCACTTCAACAAAGTCTAAG 6696
 QY 2075 CysProProGlnGlnIlyGlyPheProAlaThrProPheGlyAlaGlnIlyPserGlyThr 2094

Db 6697 TGCCTCCACAGCAGATAGCTTTCCAGCTACCCCATTTGGCGCTCAATGAGTGGAGC 6756
 QY 2095 GlyGlyProAlaProGlnProLeuGlyGlnPheGlnProValGlyThrAlaSerLeuGln 2114
 Db 6757 GGTGGCCCAAGCACACAGCAGCTTGGCCAGTTCCTGAGGAGTGGTCTCTTGGCAG 6816
 QY 2115 AsnPheAsnIleSerAsnLeuGlnLysSerIleSerAsnProProGlySerAsnLeuArg 2134
 Db 6817 AATTCAATCAGCAATTTGGAGAAATCCATCAGCAACCCCAAGCTCCAACTGCGG 6876
 QY 2135 ThrThr 2136
 Db 6877 ACCACT 6882

Search completed: September 25, 2004, 05:23:57
 Job time : 1804 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: September 25, 2004, 04:08:33 ; Search time 7148 Seconds

(without alignments)
8923.562 Million cell updates/sec

Title: US-10-010-720-14

Perfect score: 10812

Sequence: 1 MSGAAEKKQSSIRPGSLFLSP.....NISLQKSIISNPPGSLNRTT 2136

Scoring table:

BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODE=frame+g2n.model -DEV=x1h
-Q=/cgrt_1/USPTO.spool/US10010720/runat.23092004.154320.25255/app.query.fasta_1.2311
-DB=EST -QPM=fastap -SUFFIX=est -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0
-UNITS=bites -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -DIST=45
-LOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTPM=pic -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10010720 -CCGN 1.1 5750 -@runat.23092004.154320.25255 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:
1: em_estba:*
2: em_esthum:*
3: em_estlin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hnc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hnc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rnd:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
C 1	2532	23.4	1537	11	BC038700
2	1909.5	17.7	4705	11	AT321328
3	1814	16.8	3490	11	BC037114
4	1658.5	15.3	2606	11	AK032505
5	1594	14.7	1089	11	BM455039
6	1565.5	14.5	3623	11	AK052468
7	1503	13.9	912	12	BT829792
8	1467.5	13.6	991	13	BX401386
9	1428	13.2	1022	12	BM904896
10	1380.5	12.8	936	13	BQ988524
11	1356	12.5	941	13	BQ956866
12	1355.5	12.5	873	13	BQ213010
13	1336	12.4	825	9	AU130018
14	1328	12.3	829	9	AU131632
15	1321	12.2	977	13	BQ058189
16	1292.5	12.0	842	14	CD244300
17	1277	11.8	805	14	CD644052
18	1274	11.8	967	13	BU145705
19	1270.5	11.8	918	13	BX415698
20	1269	11.7	845	12	BG976502
21	1260	11.7	907	29	CG499521
22	1254.5	11.6	830	9	AU132809
23	1229	11.4	830	14	CD655395
24	1219	11.3	760	13	BX871688
25	1209	11.2	721	12	BG378924
26	1207	11.2	846	14	CFE94281
27	1203	11.1	894	13	BQ117639
28	1199	11.1	751	14	CFE53228
29	1197	11.1	905	14	CFE53228
30	1196	11.1	836	14	CD655715
31	1180	10.9	966	12	BT103160
32	1175	10.9	894	13	BU178645
33	1167.5	10.8	1126	13	BU178645
34	1152	10.6	915	12	BG166916
35	1148.5	10.6	915	12	BG173180
36	1142	10.6	701	13	BQ781019
37	1137	10.5	698	12	BG722317
38	1120	10.4	697	14	CF727167
39	1117.5	10.3	773	14	CD802866
40	1115	10.3	676	12	BG768128
41	1113	10.3	731	14	CF737595
42	1108.5	10.3	771	9	AA915514
43	1103.5	10.2	858	9	AL667586
44	1103.5	10.2	894	14	CA513151
45	1091	10.1	750	14	CA511386

ALIGNMENTS

RESULT 1
BC038700/c 1537 bp mRNA linear HTC 21-OCT-2002
LOCUS BC038700 Mus musculus, clone IMAGE:1382568, mRNA.
DEFINITION BC038700
ACCESSION BC038700
VERSION BC038700.1 GI:24182477
KEYWORDS
SOURCE
MUS MUSCULUS (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 1537)
Strausberg,R.

QY 265 helysgluinalaqlueteuysglyleuenglnhisproasnillevalargpheytra 285
 Db 1167 TCTGTGAGAGAGTGTGATGCTCAAGGGGCTGCAGACCCCAACATGCTCCGCTTCTATG 1226
 QY 285 spsertrgluserthrvallysglyslsCysillevalleuvalthrquleuuewt 305
 Db 1227 ACTCTGGAAGTCGGGTGCTGAGGGGTGAGGTTTGATGCTGTGCTCAACAACATCATGA 1286
 QY 305 hrserglYthrleuYthrYthrleuYsArpHeuValMetYsilleYsValleu 325
 Db 1287 CCTGGGGACAGCTCAAGACGTAACCTGAGAGATTCGGAATATGAAGCCCGAGTCTTC 1346
 QY 325 rgsertrpCysArgGlnilleuYsglyleuGlnPheLeuHisThrArgThrProProI 345
 Db 1347 AGCGCTGAGTGGCCAGATCCTACGGGACTTCTTCTACATTCGCGAGTGCCTGCA 1406
 QY 345 leileHisArgAspLeuLysCysAspAsnillePheIleThrGlyProThrGlySerValL 365
 Db 1407 TCTCGACCCGAGATCTCAAAATGTGACAAATGCTTCATTAACGGCCCTTCAGGCTCTGTCA 1466
 QY 365 ysilleglYAspLeuGlYleuAlaThrleuYsArgAlaSerPheAlaLysSerValilleG 385
 Db 1467 AAATCGGAGATCTCGACTGGCCACGCTCAAGCCGCTCTTTCCTCAAGAGCTCATGTG 1526
 QY 385 lythrProGluPheMetAlaProGluMetYrGlnIuYsYrAspGluSerValAspV 405
 Db 1527 GGACCCCGGAATTCATGCCCCCTGAGATGTATGAGAAAGTACGACGAGCGCTGTGATG 1586
 QY 405 alYrAlaPheGluMetCysMetleuGluMetAlaThrSerGluYrProYrSerGluC 425
 Db 1587 TGTACGCTTTGGCATGTGATGCTGAGATGTGCTACGCTGAGATACCTTACTCTGAGT 1646
 QY 425 ysglnasnalaalaglnilleYrArgArgValThrSerGlyValLysProAlaSerPheA 445
 Db 1647 GTCAGAACCGACGACAAATCTACGGAAGGCTCACTTGGGACAAAGCCCAACAGCTTCT 1706
 QY 445 splyserValAlaIleProGluValLysGluIleIleGlnIuYsYrIleArgGlnAsnLysA 465
 Db 1707 ACAAAAGTGAAGATCCGACAGATGAGAGATCATTTGMAAGCTGCATCCGACGATAGA 1766
 QY 465 spGluArgYrSerIleYsAspLeuLeuAsnHisAlaPhePheGlnIuGluThrGlyY 485
 Db 1767 ATAGAGGTTCACTATCCAGATCTTCTGACCAAGGTTCTTCGAGAGAACTGTGTG 1826
 QY 485 alArgValGluLeuAlaGluLysAspAspGlyGluYsIleAlaIleLysLeuTriPheA 505
 Db 1827 TGCATGTGAGCTGGCAGAGAGAGATGATGAGAGAACCGGCTTCMAACTTGGGTGC 1886
 QY 505 rglleGluAspIleYsLysLeuYsGlyYsYrYrLysAspAsnGluAlaIleGluPheS 525
 Db 1887 GATGTGAGAGACGACACGACG---GGGGGGGGCCCAAGGACACAGCCATCGACTTC 1943
 QY 525 eepHeAspLeuGluLysAspValProGluAspValAlaGlnIuMetValGluSerGlyY 545
 Db 1944 TGTTCAGCTCGGTCCGAGTGCACGCGAAGGTGTGCTCAGAGAGATGAGACCTTGGGTCT 2003
 QY 545 yrValCysGluGluYsAspHisYsThrMetAlaLysAlaIleLysAspArgValSerLeuI 565
 Db 2004 TAGTCTGCAGAGACGACTACGACCGACGTCGCTGCTGCTGCAAGAACGGGTTCGTCGA 2063
 QY 565 leLysArgLysArgGluGlnIuGlnLeuValArgGluGluGlnIuYsLysGlnG 585
 Db 2064 TCCAGCGGAAGCGGAGAGCTAGAGAAAAGCCAGGGAA----- 2101
 QY 585 luGluSerSerLeuYsGlnIuValGluGlnSerSerAlaSerGlnThrGlyIleYsG 605
 Db 2102 -----TTGAGAG 2108
 QY 605 luLeuProSerAlaSerThrGlyIlePro---ThrAlaSerThrThrSerAlaSerValS 624
 Db 2109 TTCTCCCAACCAAGACTCAGAGACTCTCTCCAGCACTGTGTATGACTCTGCTGCCCCCA 2168
 QY 624 erThr-----GlnValGluProGluGluProGluAlaAspGlnHisGlnIuGlnIuYrG 643

Db 2169 GTCCCTCCCCCAGAGCCCGAGAGCTCAGAGCTGAGACACCACTTCCTCTCTCC 2228
 QY 643 luGlnProSerIle---SerValLeuSerAspGlyThrValAspSerGlyGlnIuYsers 662
 Db 2229 GCCATGCCAGCTACTCATTCACATCATGTGATTCGAGACTGATGCTACCTCAGCTCTCT 2288
 QY 662 eValPheThrGluSerThrArgValSerSerGlnIuThrValSerTyrglySerGlnHisG 682
 Db 2289 CCGGCTTCCTGAGATCCCTCAGACCCCTGCTTCAGCCC----- 2326
 QY 682 luGlnAlaHisSerThrGlyThrValProGluYHisIleProSerThrValGlnAlaGlnS 702
 Db 2327 -----CCGGGGGAGTCCATTC-----AGCCCCG 2351
 QY 702 erglnProHisglYvalYrProProSerSerValAlaGlnIuYsGlnIuGlnP 722
 Db 2352 CTGAGCCCCAT---CTGTGCTGCTCCCTCGGTTTTCCTGTCTGCTCAACGCTGAGCC 2408
 QY 722 roSerSerSerSerLeuThrGlyValSerSerSerGlnProIleGlnHisProGlnIuG 742
 Db 2409 CTGGCAGT----- 2416
 QY 742 luGlnIuYIleGlnIuThrAlaProProGlnIuThrValGlnIuYrSerLeuSerGlnT 762
 Db 2417 -----GACTTTCCTCTGGGACATGATGTTTCTGCTTTCAGTCAGATGCG 2462
 QY 762 hrSerThrSerSerGluAlaThrThrAlaGlnProValSerGlnProGlnAlaProGlnV 782
 Db 2463 AGGCTACG-----CCTGAGA 2477
 QY 782 alLeuProGlnValSerAlaGlyLysGlnSerThrGlnIuYsValSerGlnValAlaProA 802
 Db 2478 TGCAGCTCTAG----- 2488
 QY 802 laGluProValAlaValAlaGlnProGlnAlaThrGlnProThrThrLeuAlaSerServ 822
 Db 2488 ----- 2488
 QY 822 alAspSerAlaHisSerAspValAlaSerGlyMetSerAspGlyAsnGluAsnValProS 842
 Db 2489 -----GCTCTCAGTGCATGACGA----- 2506
 QY 842 eSerSerGlyArgHisGluGlyArgThrThrLysArgHisYrYrArgLysSerValArgS 862
 Db 2507 -----GAAAGGGGGCAATGAGAAATCCAGTGAAGACTCTCGGAC 2549
 QY 862 eArgSerArgHisGluYsThrSerArgProLysLeuArgIleLeuAsnValSerAsnL 882
 Db 2550 GGAGA-----CCTCGATCCCGGCTTCGGGTCCAGACGATGCTCAGACC 2591
 QY 882 ysgLYAspArgValAlaGluCysGlnLeuGluThrHisAsnArgLysMetValThrPheL 902
 Db 2592 AGAGGCACAGATCGTGGAGTGCACGCTCAAACTCAACAAGCAAGATGCTACCTTTC 2651
 QY 902 yspPheAsnLeuAspGlyAspAsnProGluGlnIleAlaThrIleMetValAsnAsnAsp 922
 Db 2652 GATTTGATCTGGATGGAGACACGCCCAAGAAATGCACACTGCAAGTGTATTAATGAAT 2711
 QY 922 heIleLeuAlaIleGluArgGluSerPheValAspGlnValArgGluIleIleGluLysA 942
 Db 2712 TCATCTGCTCCCTCGAGAGCAGACGATTCCTAGCCGATTCGGGAGATTATCCAGCGAG 2771
 QY 942 laAspGlnMetLeu-----SerGluAspValSerValG 953
 Db 2772 TGGAGACCTGCTGTAAGAGAGATGCTGTGCTCGAGAGCCACTGAGATGCGTGAAGCC 2831
 QY 953 luProGluGlyAspGlnIuYsGlnIuSerLeuGlnIuYsAspAspYrGlyPheSerG 973
 Db 2832 CTCAGAGAGACCAAGCAACCAAGCTGCTCCAGGCCCATTCCTCAT-----GAGG 2882
 QY 973 lySerGlnLysLeuGluGluPhe-----LysGlnProIle----- 985

```

Db      2883 GAAACCAGCTGGTGGCCATTCACAGTCACTTCACAGGAAACAGAGAGCCCTC 2942
Qy      986 -----Proh1aser-----SermerProglInglInleGly1LeProthrserLeut 1002
Db      2943 CCTACACAGCTCCACACTCTCCAGACTCCCTAAAGCTTCACACCCCTCAGCTCA 3002
Qy      1002 hrgInvalValH1aserAglYargArgPhe1leValserProvalProgluserArg 1021
Db      3003 CCTCA-----GAGACTTACAGACAGAGAGAGCTCTCAGAGAGACAGAGACAGG 3056
Qy      1022 -----LeuArgInserLyserValPheProserGlu1leThrAspThrValAla1aserT 1040
Db      3057 AGGCTCTGGCAGAGACTGACCCGTGACCGAGGAGCTTG-----GGGCTGCGCATCGANG 3110
Qy      1040 hrgAlaInserProglY----- 1045
Db      3111 AGGAAAGATGAAAGGAAAGAAACCCCAATTGGGGGAGTTCCCAATCTTGAGCCAG 3170
Qy      1046 -----MetAsnLeuSerHisSerAlaserLeuSerLeuSerLeuGlnAlaP 1061
Db      3171 CAGCCCAAGTGTGATGAATCTACTCTTACAGACGCTGTGTCTAACAGTGAAGAGTCA 3230
Qy      1061 heser-----GluLeuArgArgAlaGlnM 1069
Db      3231 AGAGCAGCGGAGAGAGAGAGTCTGGGCTGAGCTGACAGAACTTCCGAGAGACACT 3290
Qy      1069 ethrGluGlyProAsnThr----- 1075
Db      3291 TGTGGAAGT-GGAGCACTACAGACGCTACAGAAAGAGAAATCGAAGACTTGACAGC 3349
Qy      1075 ----- 1075
Db      3350 CGGCTTGAAAGAACACCCCACTGATTTGTAGCTCCAGCGGCTATGCTCTGCGCCG 3409
Qy      1076 --AlaProPro-----AsnPheserHisThrGlyProThrPheProValVal---- 1090
Db      3410 CAGGCGCGCTCTCCAGAGAGAGCTTCCCACTCCCGCGCAACAGCTGACAGCGCTCT 3469
Qy      1091 -----ProProPheLeuSerSer1leAglValProThrThrAlaAla1ThrAlaP 1109
Db      3470 GATCTCCCTG-GCCCTGTGACATGACGAGAA-----ACTCTGTGAGTGGCGACACAC 3522
Qy      1109 rovalProAlaThrSerSer-----ProProAsnAsp1leSerThrS 1123
Db      3523 CGGCTCCAGAGAGAGCGGCGCAAGAGGGGCTGACATTCGCCGGG-GAGTGTGCGACGA 3581
Qy      1123 exVal1leGInserGluValThrValProThrgInglInleAglValAla1ThrS 1143
Db      3582 TGGTTCGAGCGGAGAGCAAACTTATAGAAAGAGAGAGAAATCCAGAGTGTGACAGAG 3641
Qy      1143 ethrGly 1145
Db      3642 ATTCTGGA 3649

RESULT 3
BC037114 3490 bp mRNA linear HTC 19-NOV-2003
LOCUS Mus musculus protein kinase, lysine deficient 4, mRNA (cdna clone
DEFINITION IMAGE:4507276), with apparent retained intron.
ACCESSION BC037114
VERSION BC037114.1 GI:23331124
KEYWORDS HTC.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 3490)
AUTHORS Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bat,C.N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
Ditchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,

```

TITLE

JOURNAL MEDLINE

PUBMED

AUTHORS

REMARK

COMMENT

NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: The Cepko Laboratory
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)
 DNA Sequencing by: National Institutes of Health Intramural
 Sequencing Center (NISC),
 Gaithersburg, Maryland;
 Web site: <http://www.nisc.nih.gov/>
 Contact: nisc_mgc@nhgri.nih.gov
 Akhter,N., Ayele,K., Beckstrom-Stenberg,S.M., Benjamin,B.,
 Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S.,
 Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P.,
 Hansen,N., Ho,S.-L., Karlins,E., Khong,P., Laric,P., Legaspi,R.,
 Maduro,O.L., Masiello,C., Maskeri,B., Mastrian,S.D., McCloskey,J.C.,
 McDowell,J., Pearson,R., Stantipop,S., Thomas,P.J., Touchman,J.W.,
 Tsurgoun,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,L.,
 Young,A., Zhang,L.-H. and Green,B.D.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LIML at: <http://image.liml.gov>
 Series: IRAC Plate: 81 Row: b Column: 5
 This clone has the following problem: retained intron.

FEATURES

source

```

1. 3490
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:4507276"
/tissue_type="Eye, retina, mouse strain C57Bl/6"
/clone_id="NIH_MGC_94"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"

```

ORIGIN

Alignment Scores:

```

Pred. No.: 1.2e-81 Length: 3490
Score: 1814.00 Matches: 521
Percent Similarity: 48.13% Conservative: 160
Best Local Similarity: 36.82% Mismatches: 357
Query Match: 16.78% Indels: 381
DB: 11 Gaps: 44

```

US-10-010-720-14 (1-2136) x BC037114 (1-3490)

```

Qy      9 GInserethrProglYserLeuPheLeuSerProProAlaProAlaProLyAsnGly 28
Db      61 CAGGAGCACTCCAGCTCTCTGCTTTCCGTTCTGCTGCGGCGCTGCGCTTCCCACTCT 120

```

```

QY 29 SerSerSerAspSerSerValGlyGluLysLeuGlyAlaAlaAlaAlaAlaValThr 48
   |||||
Db 121 AGCACCTCG-----AAATACGAGAGCTGGGGCTCC-----150

QY 49 GlyArgThrGluGluTyrArgArgArgArgIsthMetAspLysAspSerArgGlyAla 68
   |||||
Db 151 -----CATGCCAAACGAGGCTGACCTGGCTCT 180

QY 69 AlaAlaThrThrThrThrThrGluHisArgPheArgAspSerValIleCysAspSer 88
   |||||
Db 181 GCGGCC-----ATCTCCGGCTCT 198

QY 89 AsnAlaThrAlaLeuGluLeuProGlyLeuProLeuSerLeuProGlnProSerIlePro 108
   |||||
Db 199 CACCTTACGGGGCCGACCGCTGGGGCCACCTCTCG-----CCAGAGCGCCG 249

QY 109 -----AlaAlaValProGlnSerAlaProProGluProHis 120
   |||||
Db 250 CTCTCTGGGAGAGGCTGAGCTCGCGCGCTTTCGAGACCTGACCGCGCTCAGT 309

QY 121 ArgGluGluThrValThrAlaThrAlaThrSerGlnValAlaGlnGlnProProAlaAla 140
   |||||
Db 310 CGA-----TCTGGACTGCTGAGCTCTGCTGCTCAACACCTCACTCTCTCC 357

QY 141 AlaAlaProGlyGluGlnAlaVal-AlaGlyProAlaProSerThrValProSerSerTh 160
   |||||
Db 358 GGAACCCCGGATCCTCCAGACTCCGCTGGCCCCACAGAGAGC-----CCACTTCAAG 411

QY 160 rSerLysAspArgProValSerGln-----168
   |||||
Db 412 CTCTAAAGAACCCCGCAGGAGCAGATGATGGGGCAGCTCCCGTGAAGGCTGAGACT 471

QY 169 -----ProSerLeuValGlySerLysGluGluProProAlaArgSerGlySerG 186
   |||||
Db 472 TGCATGCTCGAGCTTACGGAATCT-----TCAGG 501

QY 186 YGlyGlySerAlaLysGluPro-----GlnGluGluLysArgSerG 199
   |||||
Db 502 AGGTTCAGAGGTCCCGGAGCCACCGAGGCTTCTGATGCTGACCGCCGGAGACGCGCG 561

QY 199 ngInGlnAspArgIleGluGluLeuGluThrLysAlaValGlyMetSerAsnAspGlyArg 219
   |||||
Db 562 GGGAGCGAGAAAGAAAGACACAGAACCCAGGCTGGCCAACTCTCCGACGCGCG 621

QY 219 gPheLeuLysPheAspIleGluIleGlyArgGlySerPheLysThrValTyrIleGlyLe 239
   |||||
Db 622 ATACCTCAAGTTGACATTAAGATTGAAAGAGTTCCTCAAGACCGTGTATCAGAGGCT 681

QY 239 LAspThrGluThrThrValGluValAlaTyrCysGluLeuGlnAspArgLysLeuThrI 259
   |||||
Db 682 AGACACCGACACACACGAGTAGAGTGGCTGTGTGAGTGCAGACCGGAAATCTCTCTG 741

QY 259 sSerGlnAspArgGlnArgPheLysGlnGluAlaGluMetLeuLysGlyLeuGlnHisProAs 279
   |||||
Db 742 GGGCGAGCGCGAGCGGTTCTTCGAGGAAGTTGAGATCTCAAGGGGCTCAGACCCCAA 801

QY 279 nIleValAlaPheLysAspSerTrpLysSerThrValLysGly-LysLysCysIleVal 299
   |||||
Db 802 CATCGCGCTCTTAAGACTCTCGAAGAGTGGCTGAGGGGCGAGG-----848

QY 299 euValThrGluLeuMetThrSerGlyThrLeuLysThrTyrLeuLysArgPheLysVal 319
   |||||
Db 849 -----TACCTGAGGCGGTTCCGGAAG 870

QY 319 eLysIleLysValLeuArgSerTrpCysArgGlnIleLeuLysGlyLeuGlnPheLeuH 339
   |||||
Db 871 TGAAGCGCCGAGATCTTCAGCGCTGAGTGGCCAGATTCTACGGGAGATTCTATCTTC 930

QY 339 iSthArgThrProProIleIleHisArgAspLeuLysCysAspAsnIlePheIleThng 359
   |||||
Db 991 ATTCCGAGGTGCCCTCCTGACACCGAGATCTCAAGTGTGACAAACGTTCTTATTA 990

```

```

QY 359 lYProThrGlySerValLysIleGlyAspLeuGlyLeuAlaThrLeuLysArgAlaSerP 379
   |||||
Db 991 GCGCTTCAAGGCTCTGTCAAAATCGAGACTCGGACTGGCCACGCTCAAGCGCGCTCTCT 1050

QY 379 heaLalysSerValIleGlyThrProGluPheMetAlaProGluMetTrpTrpGluLysT 399
   |||||
Db 1051 TTGCCAAGAGGTATATGGAGCCCGGAAATTCATAGGCTCTGAGATGTACAGAGAAAGT 1110

QY 399 YrAspGluSerValAspValTyrAlaPheGlyMetCysMetLeuGluMetAlaThrSerG 419
   |||||
Db 1111 ACAGACGAGGCTGTGATGTGTACGCTTGGCATGTGATCTGAGATGGCTACGTCTG 1170

QY 419 luyTrProTyrSerGluCysGlnAlaAlaGlnIleTyrArgArgValThrSerGlyV 439
   |||||
Db 1171 AGATTCCTCTCTCGATGTGATGATCAGACAAATCTTACCGAAGTCACTTCGGGCA 1230

QY 439 aLysProAlaSerPheAspLysValAlaIleProGluValLysGluIleLeuGluGlyC 459
   |||||
Db 1231 CAAGCCCAACAGCTTCTACAGAGTGAAGATGCCAGAGTGAAGAGATCATTAAGGCT 1290

QY 459 YsIleArgGlnAsnLysAspGluArgTyrSerIleLysAspLeuLeuAsnHisAlaPheP 479
   |||||
Db 1291 GCATCCGACAGGATTAAGAAAGAGAGGTTCAACATCCAGATCTCTGGCCCATGCTTCT 1350

QY 479 heGlnGluGluThrGlyValArgValGluLeuAlaGluGluAspArgGlyGluLysIleA 499
   |||||
Db 1351 TCCGCGAGGAGACAGAGGTGTGACGTGAGCTGGACGAGAAATGATGAGAGAGACCGG 1410

QY 499 lalIleLysLeuTrpLeuArgIleGluAspIleLysLysLeuLysGlyLysTyrLysAsp 519
   |||||
Db 1411 GCCTCAAGCTCTGCTGCTGCGATGAGATGCGGGCGC-----GGGGGCGCTCCACAGACA 1467

QY 519 sngLualalIleGluPheSerPheAspLeuGluLysArgAspValProGluAspValAlaGln 539
   |||||
Db 1468 ACCAGGCATCGAGTTCCTGTTCACTCGCGCGGAGCGCTGGGAGAGAGTGGCTCAGG 1527

QY 539 lmeTValGluSerGlyTyrValCysGluGlyAspHisLysThrMetAlaLysAlaIleL 559
   |||||
Db 1528 AGATGGTAGCTTGGGCTTAGTGTGAAAGTACTACACAGCCAGTGGCCCGCAGATTCT 1587

QY 559 YsAspArgValSerLeuLysArgLysArgGlnGlnGlnLeuValArgGlnGluG 579
   |||||
Db 1588 GAGAACGGGTGTGCTATTCAGCGGAGACGGAGAAAGCTTCAAAAGCTAGAGAA---1643

QY 579 lngLulysLysGlnGluGluSerSerLeuLysGlnGlnValGluGlnSerSerAlas 599
   |||||
Db 1643 -----1643

QY 599 eGlnThrGlyLysGlnLeuProSerAlaSerThrGlyLeuProThrAlaSerThT 619
   |||||
Db 1644 -----TTGAGGTTCTCCACCAAGACTCA---GAACTCTCCAGCAACGTGT 1689

QY 619 hSerAlaSerValSerThrGlnVal-----GluProGluGluProGluAlaAspG 636
   |||||
Db 1690 CTGTGGCTCGGGGTCCCCCAGATGCTTCCCCCAGAGCTGAGAGGACGAGAGGCTGAGC 1749

QY 636 lnhIleGlnLeuGlnTyrGlnGlnProSerIle---SerValLeuSerAspGlyThrV 655
   |||||
Db 1750 AGCATCAATCTTCTCTCCGATGCCAGTCACTACATCAATCATCTGATTGGGAGA 1809

QY 655 aLAspSerGlyGlnGlySerSerValPheThrGluSerArgValSerSerGlnGlnThV 675
   |||||
Db 1810 CTGATGGCTACTCAGTTCTCTCGGCTTCTGGAATCCAGACCTGCCCTTCAGCCCC- 1868

QY 675 aLserTyrGlySerGlnHisGlnGlnAlaHisSerThrGlyThrValProGlyHisIleP 695
   |||||
Db 1869 -----CTGGGGGGGTTAC 1881

QY 695 rOserThrValGlnAlaGlnSerGlnProHisGlyValTyrProProSerSerValAlaG 715
   |||||
Db 1882 CATCCAG-----1889

QY 715 lngLysGlnSerGlnGlnGlnProSerSerSerSerLeuThrGlyValSerSerSerGlnP 735

```


Percent Similarity:	96.69%	Conservative:	0
Best Local Similarity:	96.69%	Mismatches:	8
Query Match:	14.74%	Indels:	4
DB:	12	Gaps:	0
US-10-010-720-14 (1-2136) x BM455039 (1-1089)			
QY	1808	PheAsnSerSerTyMetSerSerAspAsnGluSerAspGluAspLeuLys	1827
DB	2	TTTAACTCTCTTACATAGTAGGACCAATGAGTATTCGAAGATGAAGACTTAAG	61
QY	1828	LeuGluLeuAArgLeuAArgAspLysHisLeuLysGluLeuGlnAspLeuGlnSerArg	1847
DB	62	TTAAGGCTGCAGACACTAAGAGATTAACATCTCAAGAGATTCAAGACTGCAAGTGC	121
QY	1848	GlnLysHisGlnIleGluSerLeuTyThrLysLeuGlyLysValProProAlaValIle	1867
DB	122	CAGAAAGCATGAATTTGAATCTTTGTATTCGAAGCTGGGCAAGGAGGCCCTGCTGTATT	181
QY	1868	ILeProProAlaAlaProLeuSerGlyAArgAArgArgProThrLysSerLysGlySer	1887
DB	182	ATTCCCTCCAGCTCTCCCTTTGAGGAGAAAGACGAGCCACTAAAGCAAGGAC	241
QY	1888	LysSerSerArgSerSerSerSerLeuGlyAsnLysSerProGlnLeuSerGlyAsnLeuSer	1907
DB	242	AAATCTAGTGAAGACAGTCTCTTGGGGAATTAAGCCCTTCAGGTAACCTGTCT	301
QY	1908	GlyValSerAlaAlaSerValLeuHisProGlnGlnThrLeuHisProProGlyValIle	1927
DB	302	GGTCAGAGTCAAGTTCAGTCTTGCACCCCGACAGACCTTCACCTCTCGGCAATC	361
QY	1928	ProGluSerGlyValAsnGlnLeuLeuGlnProLeuLysProSerProSerSerAspAsn	1947
DB	362	CCAGAGTCCCGGAGATCAGCTGTGTACAGCCCTTAAGCATCTCCCTCAGTGAC	421
QY	1948	LeuTySerAlaIleThrSerAspGlyValIleSerValProSerLeuSerAlaProGly	1967
DB	422	CTTATATTCACCTTCACAGTGAATGGTCCATTCATTCACCAAGCCCTTCTGCTCCAG	481
QY	1968	GlnGlyThrSerSerThrAsnThrValGlyAlaThrValAsnSerGlnAlaGlnAla	1987
DB	482	CAAGGAACCGACGACCAACACTGTTGGGGCAACGATGAACGCAAGCCGCGCAAGCT	541
QY	1988	GlnProProAlaMetThrSerSerArgLysGlyThrPheThrAspAspLeuHisLysLeu	2007
DB	542	CAGCCCTCCGTCATGAGTCCACAGGAGGACATTCACATACATCTTGCAAGATTG	601
QY	2008	ValAspAsnTrpAlaArgAspAlaMetAsnLeuSerGlyAArgArgLysGlyHis	2027
DB	602	GTAAGCAATTTGGGCGGAGATGCAATGATCTCTCAGGAGGAGAGCAAGCAAGGGCAC	661
QY	2028	MetAsnTyGlnGlyProGlyMetAlaArgLysPheSerAlaProGlyGlnLeuCysIle	2047
DB	662	ATGAATTAAGAGGCGCTGGAAGGCAAGGAGATTCTTGACACTTGAGCAACTGTTCATC	721
QY	2048	SerMetThrSerAsnLeuGlyLysSerAlaProIleSerAlaAlaSerAlaThrSerLeu	2067
DB	722	TCCATGACCTCGAAGCTGGTGGCTGCGCCCATCTTGACGACATCACTACCTCTCTA	781
QY	2068	GlyHisSphThrLysSerMetCysProProGlnGlnTyGlyPheProAlaThrProPhe	2087
DB	782	GGTCACTTACCAAGTCTATGTGCCCCCAAGCAGATATGGCTTCCACTAC-CGATT	840
QY	2088	GlyAlaGlnTrp-SerGlyThrGlyValProAlaProGln-ProLeuGlyGlnPheGlnP	2107
DB	841	GGCGGCTCATAGGAGTGGGAGCCAGTGGCCCAACACAGCGCCCACTTGCGCAGTTTCAA	900
QY	2107	roValGlyThrAlaSerLeuGlnAsnAsnIleSerAsnLeuGlnLysSerLysSerA	2127
DB	901	CTGTGGGAGTGGCTCTTGCAAGAAATTTCAACATCAGCAATTTGCAAGAAATCCATCAGCA	960
QY	2127	sn-ProProGlySerAsnLeuArgThr 2136	

DB	961	ACCCCCCAGGCTTCAACTGCGGACACC	990	
RESULT 6	AK052468			
LOCUS	AK052468			
DEFINITION	AK052468	3623 bp	mRNA	
ACCESSION	AK052468			
VERSION	AK052468.1	GI:26095199		
KEYWORDS	HTC; CAP trapper.			
SOURCE	Mus musculus (house mouse)			
ORGANISM	Mus musculus			
REFERENCE	1	Carninci, P. and Hayashizaki, Y.	linear HTC 20-SEP-2001	
AUTHORS				
TITLE	High-efficiency full-length cDNA cloning			
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)			
MEDLINE	99279253			
PUBMED	10349636			
REFERENCE	2	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes Genome Res. 10 (10), 1617-1630 (2000)		
TITLE	3			
JOURNAL	20499374			
MEDLINE	11042159			
PUBMED				
REFERENCE	4	Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Kono, H., Akiyama, J., Nishi, K., Kitsuana, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujisake, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watabiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuda, S., Kawai, O., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y. RIKEN integrated sequence analysis (RISA) system -384-format sequencing pipeline with 384 multicapillary sequencer Genome Res. 10 (11), 1757-1771 (2000)		
TITLE	5			
JOURNAL	20530913			
MEDLINE	11076861			
PUBMED				
REFERENCE	6			
AUTHORS				
TITLE	The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.			
JOURNAL	Functional annotation of a full-length mouse cDNA collection			
REFERENCE	Nature 409, 685-690 (2001)			
AUTHORS				
TITLE	7			
JOURNAL	The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.			
REFERENCE	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs Nature 420, 563-573 (2002)			
AUTHORS	6 (bases 1 to 3623)			
TITLE	Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Kato, H., Kawai, D., Kojima, Y., Kondo, S., Kono, H., Konda, M., Koyama, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Nunazaki, R., Ono, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shingawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takeuchi-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.			
TITLE	Submitted (16-JUL-2001)			
JOURNAL	Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsukumi-ku, Yokohama, Kanagawa 230-0045, Japan [E-mail: genome-res@gsc.riken.go.jp]			


```

http://image.llnl.gov
Plate: L1AM11426 row: n column: 15
High quality sequence stop: 899.
Location/Qualifiers
1. 912
FEATURES
source
```

ORIGIN

/clone_lib="NIH_MGC_119",
/name="Organ: brain; Vector: pCMV-Sport6; Site_1: NCI,
Site_2: ECoRV (destroyed); RNA source normal medulla from
anonymous male age 27. Library is oligo-dT primed and
directionally cloned. ECoRV site is destroyed upon
0.9-1.3 kb. Average insert size 1.3 kb, insert size range
full-length clones and was enriched for
(Invitrogen). Research Genetics tracking code 013. Note:
this is a NIH_MGC Library."

US-10-010-720-14 (1-2136) X BI829792 (1-912)

QY	1821	IleGlnAspLeuSpIusSpleuLysLeuGIleuHrgyrcrLeuAgaPlySH:LSeuLYSGlu	184
Dd	1	ATCGAAGATGAAGAAGCTTAAGATTGAGCTGGACACTACAGATATAAACATCTTAAAGAG	60
QY	1841	IleGlnAspLeuGIuSerArGInLYSHISGLurIeGIuSerLeuTYrThLYSLeuGIy	186
Dd	61	ATTCCAGAGCACTGCAGAGTCGCCCAAGAACATGAATTGATCTTGIMTACCMAACTGGGC	120
QY	1861	LysValProProAlaValIlelleProProAlaAlaProLeuSerGIyaGaYaArgArG	188
Dd	121	AAGGTGCCCCCTCCTGTTATTATTATTCCTCCAGCTGCTCCCTTCAGGAGAGAACGACGA	180
QY	1881	ProThrLyseSerLySGLyserLYSSerSerArGserSerSerLeuGIyaAnLYSserPro	190
Dd	181	CCCACATAAAAGCAAAGSCAGCAATCTAAGTCGAAGCATTCCTGGGGAATTAAGGCCCC	240
QY	1901	GlnLeuSerGIyaSnLeuSerGIyGInSerAlaAlaSerValIeuHiSproGInGInThr	1920
Dd	241	CAGCTTTCAGGTAACTGTCTGTGCAGAGTCGAGCTTCAGTCTTGACCCCCAGCACACC	300
QY	1921	LeuHiSProProGIyaSnIeProGIuSerGIyGlaSncInLeuGInLeuGInProLeuLYS	1940
Dd	301	CTCCACCTCTCTGGCAACATCCAGAGTCCGGGCAAGATAGCTGTAAAGCCCTTAG	360
QY	1941	ProSerProSerSerAspSnLeuTYrSerAlaPheThSerAspGIalailSerVal	1960
Dd	361	CCATCTCCCTCCAGTAGCAACCTCTATTACGCTTCAACAGTATGTGTGCATTTCACTA	420
QY	1961	ProSerLeuSerAlaProGIyGInLYrThSerSerThraSnThryAlaIarVal	1980
Dd	421	CCAAGCTTTTCCTCCAGGTCAAGAACACGACAGACAACAACTGTGGGGCAACAGTG	480
QY	1981	AsnSerGlnAlaAlaGlnIaGInProProAlaMetThrSerSerArGlySGlyThrPhe	2000
Dd	481	AACACGCCAAGCCGCCAAGCTCAGCTCTCTGCCATACGTCACAGCGAAGGGCAACATTC	540
QY	2001	ThraSapLeuHiSlLyLeuValAspAsenTrpAlaArgAspAlaMetAsnLeuSerGIy	2020
Dd	541	ACAGATGACTTSCACAAGTTGGTAGCAATTGGGCCCGAGATGCCATGATCTCTCAGGC	600

Qy	2021	ATGAGAGGLASetLeuGIYHIsheRsnTyrGlUGlUcyProGIYMeLAlaIrgLyPheSer	2040
Db	601	AGGAGAGGAAGAAGGGCACTGATTATAGAGGGCCCTTGAAATGGCAAGGAACTTCCT	660
Qy	2041	AlAProGIYgInLeuCysIleSerMetThxSeranLeuGIYglYSerAlaProIleSer	2060
Db	661	GACCTGGGGCAACTGTGCATCTCCATGACCTCGAAACCTGGGTGGCTCTGCCCATCTCT	720
Qy	2061	AlaAlaSerAlaThxSerLeuGIYHIsheRhILysSerMetCysPProGIgInITyr	2080
Db	721	GCAGATCACGCTACTCTCTTAGTACTCTCAACCAAGCTAAGTCCCCCAAGCAGATAT	780
Qy	2081	GIYPheProalArThProPhelGIYAlaGIInTrpSerGIYThrGIYglYProAlaProGIIn	2100
Db	781	GGCTTTCCAGCTACCCCATTTGGGGCTCAAGAGAGTGGAGCGGTGGCCCAAGACCACA	840
Qy	2101	ProLeuGIYgInPheGIInProValGIYThrAlaSerLeuGIInaPheasnile-Seras	2120
Db	841	GCATCTGGCAG-TTTCACACCTGTGGAAACGTGCTCTTGAGATTTCAACATCAGCA	900
Qy	2120	nLeuGIInLysSer	2124
Db	900	TTTGCAAAAATCC	912
RESULT 8	BX401386	991 bp	mRNA linear EST 13-MAY-2003
LOCUS	BX401386 Homo sapiens HEILA CELLS COT 25-NORMALIZED Homo sapiens		
DEFINITION	CDNA clone CS0DK012YH04 5-PRIME, mRNA sequence.		
ACCESSION	BX401386		
VERSION	BX401386.1 GI:30631970		
KEYWORDS	EST.		
SOURCE	Homo sapiens (human)		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 991) Li,W.B., Gruber,C.J., Tessee,J. and Polayes,D. Full-length cDNA libraries and normalization Unpublished (2001) Contact: Genoscope Genoscope - Centre National de Sequenage BP 191 91006 EVRY cedex - France Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr was normalized. Library was constructed by life technologies, a division of invitrogen. This sequence belongs to sequence cluster 6600.f For more information about this cluster, see http://www.genoscope.cns.fr/ cgi-bin/cluster.cgi?seq=CS0DK012DD02QP1&c[cluster=6600.f. Contact : Feng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/ Invitrogen Corporation 1600 Faraday Avenue Genosope Sequence ID : CS0DK012DD02QP1. Location/Qualifiers 1..991 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="CS0DK012YH04" /cell_line="HELA CELLS COT 25-NORMALIZED" /cell_line="HELA" /clone_lib="Homo sapiens HEILA CELLS COT 25-NORMALIZED" /note="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."		
FEATURES	source		
ORIGIN			
Alignment Scores:			
Pred. No.:	9, 08e-65	Length:	991
Score:	1467.50	Matches:	304
Percent Similarity:	96.82%	Conservative:	0
Best Local Similarity:	96.82%	Mismatches:	10
Query Match:	13.57%	Indels:	2

Alignment Scores:	
Pred. No.:	9.08e-65
Score:	1467.50
Percent Similarity:	96.82%
Best Local Similarity:	96.82%
Query Match:	13.57%
Length:	993
Matches:	30
Conservative:	0
Mismatches:	10
Indels:	2

DB: 13 Gaps: 0

US-10-010-720-14 (1-2136) x BX401386 (1-991)

QY 1182 SerProSerLeuGlnValProThSerThSerGluValValSerSerThraLeu 1201

DB 52 TCCCGGAGATCTTCAAGTCCCAATCCATCATCTGAGTGTGTTCTTAGTACACAGC 111

QY 1202 ThyProSerValThValSerLarThSerLarValAGlyGlySerThraLarPro 1221

DB 112 TATCTCTCGTAGACAGTTTCAGCACTTACGCTTGCAGGGGAGAGTAGTCTCAACCCA 171

QY 1222 GlyProLyProProAlaValValSerGlnGlnAlaAGlySerThThraValGlyAla 1241

DB 172 GGTCTTACGCTTCCAGCTTATGATTCACAGCAGCAGCAGCAGCAGCAGCAGCAGC 231

QY 1242 ThrLeuThSerValSerThThThThSerPheProSerThraLarSerGlnLeuSerTle 1261

DB 232 ACATTATACATCAAGTTCTTACACCACTTCTATCCCAAGCAGCAGCAGCAGCAGCAGC 291

QY 1262 GlnLeuSerSerSerThSerThProThThraLeuAlaGluThThraValValSerLarHs 1281

DB 292 CAGCTTAGACAGAGTACTTCTTACTTCTTACTTACTTACTTACTTACTTACTTACTTACTT 351

QY 1282 SerLeuAspLySerThSerThSerThSerThThGlyLeuAlaPheSerLeuSerLarPro 1301

DB 352 TCCTTAGATAGACATCTCATAGCAGTACACAGTGGCTTCTCCCTCTCTGCACCA 411

QY 1302 SerSerSerSerSerProGlyAlaGlyValSerSerTyrTleSerGlnProGlyGlyLeu 1321

DB 412 TCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 471

QY 1322 HisProLeuValLleProSerValLleAlaSerThProLleuProGlnAlaAGly 1341

DB 472 CATCTTGTGTCATTCATCATGATAGATGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 531

QY 1342 ProThSerThProLeuLeuProGlnValProSerLleProProLeuValGlnProVal 1361

DB 532 CCACTTCTTACACCTTATTAATCCCAAGTACCTAGTACCTTCTTCTTCTTCTTCTTCTTCT 591

QY 1362 AlaAsnValProAlaValGlnGlnThraLleuLleHisSerGlnProGlnProAlaLeu 1381

DB 592 GCCAATGTGCTCTGCTAGACAGACACTAATCATAGTACAGCTCAACAGCAGCTTGTCT 651

QY 1382 ProAsnGlnProHsThThHsCysProGluValAlaSerSerPheThGlnProLyAlaPro 1401

DB 652 CCCAACCAGCCCACTACTTGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 711

QY 1402 GlyLleAspAspLleLysThThLeuGluGluLysLeuArgSerLeuPheSerGluHsSer 1421

DB 712 GGAATTAATGACATTAAGAAATCTTAAGAAAGCTGGGCTCTCTTCTTCTTCTTCTTCTTCTTCT 771

QY 1422 SerSerGlyAlaGlnHsAlaSerValSerLeuGlnThSerLeuValLleGluSerTh 1441

DB 772 TCATCTGAGCTCAGCATGCTCTGTCTCACTGAGACCTCACTAGTACATAGAGAGACT 831

QY 1442 ValThProGlyLleProThThraAlaValAlaProSerLysLeuLeuThSerThTh 1461

DB 832 GTCAACACAGGACCACTCACTAGCTGTGTCACCAAGCAAACTCTGAGCTTCTTCTTCTTCTTCT 891

QY 1462 SerThCysLeuProProThraLleuProLeuGlyThraValAlaLeuProValThPro 1481

DB 892 AGTACTGTGTTTCCACCAACCAATTAACCTAGGAAAGAGTGTGTTGCGAGTTACACCA 951

QY 1482 ValValThProGlyGlnValSerThProValSerThTh 1495

DB 952 GTGTCACACT-GGGCAAGTTT-TWCCMGTCAGCACTACT 991

RESULT 9

BM904896 1022 bp mRNA linear EST 12-MAR-2002

LOCUS BM904896

DEFINITION AGNCOURT 6699439 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:5557718

5', mRNA sequence.

ACCESSION BM904896

VERSION BM904896.1 GI:19355275

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

AUTHORS NIH-MGC http://mgs.nci.nih.gov/.

TITLE NIH-MGC National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D. Email: cga@bbs-remail.nih.gov Tissue Procurement: ATCC/DCTD/DTF CDNA Library Preparation: Life Technologies, Inc. CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN) DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at: http://image.llnl.gov Plate: LLM12279 row: 1 column: 15 High quality sequence start: 3 High quality sequence stop: 715. Location/Qualifiers

FEATURES

source

1..1022

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:5557718"

/issue_type="melanotic melanoma"

/lab_host="DH10B (phage-resistant)"

/clone_1b="NIH_MGC_72"

/note="Organ: skin; Vector: pCMV-SPORT6; Site: 1: NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 2 kb. Library constructed by Life Technologies."

ORIGIN

Alignment Scores:

Pred. No.: 9,54e-63 Length: 1022

Score: 1428.00 Matches: 308

Percent Similarity: 91.55% Conservative: 6

Best Local Similarity: 89.80% Mismatches: 16

Query Match: 13,214 Indels: 13

DB: 12 Gaps: 4

US-10-010-720-14 (1-2136) x BM904896 (1-1022)

QY 1302 SerSerSerSerSerProGlyAlaGlyValSerSerTyrTleSerGlnProGlyGlyLeu 1321

DB 9 TCGTCACACAGCGCTCCGGAGCAGAGTGTCTTAGTATATTTCTCAGCCTGTGGGCTG 68

QY 1322 HisProLeuValLleProSerValLleAlaSerThProLleuProGlnAlaAGly 1341

DB 69 CATCTTGTGTCATTCATAGATAGTACTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 128

QY 1342 ProThSerThProLeuLeuProGlnValProSerLleProProLeuValGlnProVal 1361

DB 129 CCACTTCTTACACCTTATTAATCCCAAGTACCTAGTATCCCACTTGTGTCAGCCTGT 188

QY 1362 AlaAsnValProAlaValGlnGlnThraLleuLleHisSerGlnProGlnProAlaLeu 1381

DB 189 GCCAATGTGCTGTGTCAGCAGACACTAATCATAGTACAGCTCAACAGCAGCTTGTCTT 248

QY 1382 ProAsnGlnProHsThThHsCysProGluValAlaSerSerPheThGlnProLyAlaPro 1401

DB 249 CCCAACCAGCCCACTACTTATTTGCTCGAAGTAGATTTCTGATACAAACCAAGCTCTCT 308

QY 1402 GlyLleAspAspLleLysThThLeuGluGluLysLeuArgSerLeuPheSerGluHsSer 1421

DB 309 GGAATTAATGACATTAAGAACTCTTAAGAAAGCTGGGCTCTCTTCTTCTTCTTCTTCTTCTTCT 368

QY 1422 SerSerGlyAlaGlnHsAlaSerValSerLeuGlnThSerLeuValLleGluSerTh 1441

```

Db      369  TCATCTGAGGTACACAGCTCTCTCTCTACAGAGACCTCACTAGTACATAGAGCAGCT 428
Qy      1442  ValThrProGlyIleProThrThraAlaAlaProSerIleLeuThrSerThrThr 1461
Db      429  GTCAACACAGGATCCCAACTACTCTCTGTTCACCAAGAACTCTCACTTCTACACCA 488
Qy      1462  SerThrCysLeuProProThrAsnLeuProLeuGlyThrValAlaLeuProValThrPro 1481
Db      489  AGTACTCTCTTACCAACCAACATTTACCACTAGGAACAGTTGCTTCCAGTTACCA 548
Qy      1482  ValValThrProGlyValValSerThrProValSerThrThrThrSerGlyValIlePro 1501
Db      549  GTGGTCAACCTGGGCAAGTTTCACTCCAGTCACTACTACTACTAGAGAGTGAACCT 608
Qy      1502  GlyThrAlaProSerIleProProLeuThrIleValAlaProValIleProValGlyThrGlu 1521
Db      609  GGAACGTGCTCTCCAGACCACTCTTACTAGAGCTCGAGTGTGCCAGTGGTACTGAA 668
Qy      1522  LeuProAlaGlyThrLeuProSerGlyGluLeuProProPheProGlyProSerLeuThr 1541
Db      669  CTTCACACAGGATCTTACCAACGAGAGAGCTGCCACCTTTCCAGAGACCTTCTAAC 728
Qy      1542  GluSerGlnGln-ProLeuGluAspLeuAspAlaGluLeuArgThrIleSerProGly 1561
Db      729  CAGTCCACGACACCTCTTACAGAGATCTGATGCTCAATGAGAAACACTTACTCTCAGA 788
Qy      1561  U***IleThrValThr-SerAlaValGlyProValSerMetAla-AlaProThrAla-I 1580
Db      789  GATGATCACAGTACCTTCTGCGGTGTCTCTGTCTCAATGCGGGCTCCACAGCCAA 848
Qy      1580  ethrGluAla-GlyThrGlnProGlnIleGlyVal---SerGlnValIleGluGlyProV 1599
Db      849  CACAAAAGCCAGGAGAACCCCTCAGAAAGTGTCTTCTCAAGTCAAGAAAGAGCCCTG 908
Qy      1599  alLeuAla-ThrSerSer-GlyValGlyValPheIleMetGlyY-----ArgPheG 1615
Db      909  TCTTACCAACTAGTTTAAAGAGTGGGGTTTTTCAGAAAGGAGCAAAATTTCCGGTTCC 968
Qy      1615  nValSerValAlaAlaAspGlyValGlnIleGlyGly---LysAsnIleSerGln 1632
Db      969  T--GGTTGACAAACAGGGGCCCAAAAGGGGTAAAGGTAATAGTCCAAA 1020

RESULT 10
LOCUS   BQ898524                936 bp    mRNA    linear    EST 16-AUG-2002
DEFINITION
AGENCOURT 8728264 NIH_MGC_47 Homo sapiens cDNA clone IMAGE:6339300
5', mRNA sequence.
ACCESSION
BQ898524
VERSION  BQ898524.1 GI:22290538
KEYWORDS
EST.
SOURCE  Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 936)
NIH-MGC http://mgi.nci.nih.gov/
Mammalian Institutes of Health, Mammalian Gene Collection (MGC)
Contact: Robert Strausberg, Ph.D.
Email: cga@bhs-neml.nih.gov
Tissue Procurement: ATCC
cDNA library Preparation: Rubin Laboratory
cDNA library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Place: LHC2533 row: 0 column: 13
High quality sequence stop: 656.
Location/Qualifiers
1..936
/organism="Homo sapiens"

```

```

/mo1_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6339300"
/issue_type="neuroblastoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 47"
/note="Organ: brain; Vector: pOT87; Site: 1: XhoI; Site: 2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGACAGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

ORIGIN
Alignment Scores:
Pred. No.:      2,196-60      Length:      936
Score:          1380.50      Matches:      293
Percent Similarity: 94.5%      Conservative: 2
Best Local Similarity: 93.91%      Mismatches: 12
Query Match:     12.77%      Indels:      5
DB:              13          Gaps:        3

US-10-010-720-14 (1-2136) x BQ898524 (1-936)
Qy      915  ThrIleMetValAsnAsnAspPheIleLeuAlaIleGluArgGluSerPheValAspGln 934
Db      1  ACAATTATGCTGAACAATGACTTATTCTAGCAATAGAGAGAGTGCTTTGTGATCA 60
Qy      935  ValArgGluIleIleGluValAlaAspGluMetLeuSerGluAspValSerValGluPro 954
Db      61  GTGCGAAGAAATTTATGAAGAAAGCTGATGAAATCTCAGTAGAGATGTCAGTGTGAACA 120
Qy      955  GluGlyAspGlnIleGluGluSerLeuGlnIleGlyAspAspPyrGlyPheSerGlySer 974
Db      121  GAGGGATGATGAGGATTTGAGAGAGCTTCAAGAGAAAGATGACTATGAGCTTTTCAAG 180
Qy      975  GlnIleLeuGlnIleGluPheIleGlnProIleProAlaSerSerMetProGlnIle 994
Db      181  CAGAAATGGAGAGAGAGATTCAAACCAATTCCTGCGCTTCATGCGACAGCAAAATA 240
Qy      995  GlyIleProThrSerSerIleThrGlnValValHisSerAlaGlyArgPheIleVal 1014
Db      241  GGGATTCCTACCAAGTCTTAACTCAAGTGTTCATTTCCGGGAGAGCGGTTTATAGTG 300
Qy      1015  SerProValProGluSerArgLeuArgGluSerIleValPheProSerGluIleThrAsp 1034
Db      301  AGTCCCTGTCACGAAAGCCGATTACAGAAATCAAAAGTTTCCCGAGTGAATTAACAGAT 360
Qy      1035  ThrValAlaIleSerThrAlaGlnSerProGlyMetAsnLeuSerHisSerAlaSerSer 1054
Db      361  ACAGTTGCTGCTCTTACAGCTCAGAGCCCTGGAATGAACCTTGTCTCACTGTGATATCC 420
Qy      1055  LeuSerLeuGlnAlaPheSerGluLeuArgAlaGlnMetThrGlnIleProAsn 1074
Db      421  CTTAGCTTCAACAGGCTTTCTGAACCTTAACGCGCCAAATGACAAAGAGACCAAC 480
Qy      1075  ThrAlaProProAsnPheSerHisThrGlyProThrPheProValValProProPheLeu 1094
Db      481  ACAGACCTCCCAAACTTTAGTATACAGACCAACATTTCCAGTAGTACTCTTTCTTA 540
Qy      1095  SerSerIleAlaGlyValProThrThrAlaAlaIleThrAlaProValProAlaThrSer 1114
Db      541  AGTACCATTTGCTGGAGTCCCAACACAGAGCAGCAGCAGCCAGTCCCTCAACAAGC 600
Qy      1115  SerProProAsnAspIleSerThrSerValIleGlnSerGluValThrValProThrGlu 1134
Db      601  AGCCCTCTTAATGACATTTCCACATCAGTAATTCGTGAGTTACAGTGCCCACTGAA 660
Qy      1135  GluGlyIleAlaGlyValAlaThrSerThrGlyValValThrSerGlyGlyLeuProIle 1154

```


Db 661 GAGGGATGCTGAGTTGCCACAGCAGAGGTGTGTAACTTCAGTGTCTCCCAATA 720

QY 1155 ProProValSerGluSerProValIleuSerSerValIleSerIleThrIleProAla 1174

Db 721 CCACCTGCTGCAATCAGCAGTACTTCCAGGAGTTCAAGATTCACAACTCTGCA 780

QY 1175 ValIleSerIleSerThrThrSer-ProSerLeuGlnValProThrSerThrSer---G1 1193

Db 781 GTTGTCTCAATATCTACTACATCCCGGCTACTTCAGTCCCAATCCACATCTGGAG 840

QY 1193 uIleValIleSerSerThrAlaIleu-TyrPro---SerValThrValSerAlaThrSera 1212

Db 841 ATCGGTTGTTCTTCAAGACAGCAGTGTATCTTCAATACAGTTTTCAGCAAACTTCAN 900

QY 1212 Ia---SerIleGlyIleSerThrAlaThrPro 1221

Db 901 NCCTCTTGAGAGGGGGGAGTACTGCTTACCC 932

RESULT 11

LOCUS BQ956866 941 bp mRNA linear EST 21-AUG-2002

DEFINITION AGENCOURT 8810997 NIH_MGC_101 Homo sapiens cDNA clone IMAGE:6429116

5', mRNA sequence.

ACCESSION BQ956866

VERSION BQ956866.1 GI:22372344

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Bkaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 941)

NIH-MGC http://mgs.nci.nih.gov/.

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgaabs-r@mail.nih.gov

Tissue Procurement: ATCC

CDNA Library Preparation: Rubin Laboratory

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LNL at:

http://image.lnl.gov

Plate: LCM2616 row: m column: 21

High quality sequence start: 4

High quality sequence stop: 623.

Location/Qualifiers

1..941

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:6429116"

/tissue_type="epidermoid carcinoma, cell line"

/lab_host="DH10B (phage-resistant)"

/clone_id="NIH MGC 101"

/note="Organ: Lung; Vector: pORF7; Site:1: EcoRI; Site:2:

XhoI; CDNA made by oligo-dt priming. Directionally cloned

into EcoRI/XhoI sites using the following 5' adaptor:

GGCAGGAG(G). Library constructed by Ling Hong in the

Laboratory of Gerald M. Rubin (University of California,

Berkeley) using ZAP-CDNA synthesis kit (Stratagene) and

Superscript II RT (Life Technologies). Note: this is a

NIH_MGC Library."

ORIGIN

Alignment Scores:

Pred. No.: 3,866-59 Length: 941

Score: 1356.00 Matches: 288

Percent Similarity: 93.65% Conservative: 6

Best Local Similarity: 91.72% Indels: 13

Query Match: 12.54% Mismatches: 8

DB: 13 Gaps: 2

US-10-010-720-14 (1-2136) x BQ956866 (1-941)

QY 1582 GluAlaGlyThrGlnProGlnIleuSerGlyValSerGlnValIleGlyProValIleuAla 1601

Db 1 GAAGGAGAAACAGACCTCAGAGAGGTGTTTCTCAAGTCAAGAGAGCCCTCTCAGCA 60

QY 1602 ThrSerSerGlyAlaGlyValPheIleuSerGlyValPheGlnValIleSerValAlaIleAsp 1621

Db 61 ACTAGTTCAGAGAGGTGTTTAAAGATGGAGCAATTCAGGTTTCTGTTCCAGCAGAC 120

QY 1622 GlyAlaGlnIleuSerGlyValPheIleuSerGlyValPheIleuSerGlyValIleSerGly 1641

Db 121 GGTGCCCAAGAAAGAGGTAATAATATGTCAGAGATGCAAACTCTCTTCAATTCATTC 180

QY 1642 SerThrSerGluSerSerValIleuSerSerSerSerProGluSerThrIleuValIlePro 1661

Db 181 AGCACTCTGAGATCTCTCAGTGTATCAAGTATGATCCAGAGATGCTTGTGTAACCA 240

QY 1662 GluProAsnGlyIleThrIleProGlyIleSerSerAspValProGluSerAlaHisIle 1681

Db 241 GAGCCGATGCAATGCAATCACTCCCTGTATCTTTCAGATGTCCAGAGAGTCCCAAA 300

QY 1682 ThrThrAlaSerGlnAlaIleuSerSerThrGlyIleProThrIleValGlyArgPheGln 1701

Db 301 ACTACTGCTCTCAGAGCAAGTCACTGAGCAGCTTCCAGAGCTTGGACCTTTTCAG 360

QY 1702 ValThrThrThrAlaAsnIleuValGlyArgPheSerValSerIleThrGluAspIle 1721

Db 361 GTGACAACTACAGCAAAAGAGGCTGTTCTCTGTATCAAACTGAGGCAAGATC 420

QY 1722 ThrAspThrIleuSerGlyIleProValAlaIleSerProPheThrAsnIleuGlnAla 1741

Db 421 ACTGACAAAGAAAGAAAGAGCAGTGCATCTCTCTTTTAAATGATTTGAAACAAGCT 480

QY 1742 ValIleuProAlaValIleProIleuSerGlyIleuSerGlyIleuSerGlyIleuSerGly 1761

Db 481 GTTCTTCTGCTGTGTATACCAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 540

QY 1762 AsnGlyProSerSerSerProGlnAlaAlaPheIleuSerArgAspValAspAspGlySer 1781

Db 541 AATGGGCGGCTTCTGATACCCGAGGCGGCTTTTAAATGATGATGATGATGATGATGATG 600

QY 1782 GlySerProHisSer-ProHisGlnIleuSerSerIleuProSerGlnIleuSer 1801

Db 601 GGTAGTTCACACTCGCCCATCAGTCACTCAAAAGCCCTTCATGCAAGATTCAG 660

QY 1801 ArgIleuSerSerSerSerPheAsnSer-SerTyrMetSerSerSerAspAsnGlySerAsp 1821

Db 661 TCAAAAGCTTAAATTCATTTTATCTCTCTTTCATGATGATGATGATGATGATGATG 720

QY 1821 IeGluAspGluAspIleuSerGlyIleuSerArgIleuSerArgIleuSerArgIleuSerArg 1841

Db 721 TCGAAGTGAAGACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 780

QY 1841 IeGln-AspIleuSerArgIleuSerGlyIleuSerArgIleuSerArgIleuSerArg 1860

Db 781 TTCAGGAGCTGAGAGATCCCGGAGCAAGAAATTTGATTCCTTTGATTCCTTCAAACTGG 840

QY 1860 ValValProPro---AlaValIleIleProProAlaAlaProIleuSerGlyArgArg 1879

Db 841 CAAAGTGGCCCGCTGCTGTTATTTATTCCTCCAGACTGGCTCTCTTTCCGGGA-AAAAG 899

QY 1879 GARGProThr-----LysSerIleGlySerIle 1888

Db 900 ACGACCAACCACTTAAGCAAGGAGGAGCA 933

RESULT 12

BQ213010/c 873 bp mRNA linear EST 02-MAY-2002

LOCUS BQ213010

DEFINITION AGENCOURT 7554697 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:6062979

5', mRNA Sequence.

ACCESSION BQ213010

VERSION BQ213010.1 GI:20393787

Alignment Scores:

Pred. No.:	Score:	Best Local Similarity:	Query Match:	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
3	38e-58	1336.00	98.91%	825	271	1	3	1	0

US-10-010-720-14 (1-2136) x AU130018 (1-825)

1642 SerThrSerGluSerSerValLeuSerSerSerProGluSerThrLeuValysPro 1661

2 AGACCTCAGAGCTCCAGTGCATCAAGAGTCCAGAGTCTGTTGAAACCA 61

1662 GluProAsnGlyTleThrIleProGlyIleSerSerAspValProGluSerAlaHisLys 1681

62 GAGCCGAATGGCATCAATCCCTGGTATCTTCCAGATGTGCCAGAGTCCCAAA 121

1682 ThrThrAlaSerGluAlaLysSerSerProGluSerValGlyArgPheGln 1701

122 ACTACGCTCAGAGCAAGCAAGTACAGACTGGGAGCTCAAGAGTTGACCTTTGAG 181

1702 ValThrThrThrAlaAsnLysValGlyArgPheSerValSerLysThrGluAspLysIle 1721

182 GTGACAACTACAGCAAAAGAGGCTCTTCTGATCAAAAAGTCCAGAGATC 241

1722 ThrAspThrLysLysGluGlyProValAlaSerProProPheMetLeuGluGlnAla 1741

242 ACTGACCAAAAGAGAGAGCAAGTGGCATCTCTCTTTATGATTTGAAACAAGCT 301

1742 ValLeuProAlaValIleProLysLysGluLysProGluLeuSerGluProSerHisLys 1761

302 GTTCTTCTCTGCTGATCAACAAAGAGAGAGCTGACCTGTCAGAGCTTACATCTA 361

1762 AsnGlyProSerSerAspProGluAlaAlaPheLeuSerArgAspValAspAspGlySer 1781

362 AATGGGCGCTCTTCTGACCCGAGAGCGCTTTTAAAGTAGGAGATGATGATGTTCC 421

1782 GlySerProHisSerProHisGlnLeuSerSerLysSerLeuProSerGlnAsnLeuSer 1801

422 GTTAGTCCACACTGCGCCCACTAGCTAGCTCAAAAGCGCTTCTTACCCAGAACTTAGT 481

1802 GlnSerLeuSerAsnSerPheAsnSerSerTyMetSerSerAspAsnGlnSerAspIle 1821

482 CAAAGCCTTAGTATTATTAATCTCTTACATGATGACGACATGATGATGATGATC 541

1822 GluAspGluAspLeuLysLeuGluLeuArgArgLeuArgAspLysHisLysGluIle 1841

542 GAAAGTAAAGACTTAAAGTTAGACTGCGACGACCTAGAGATTAACAATCTCAAGGAT 601

1842 GlnAspLeuGlnSerArgGlnLysHisGluIleGlnSerLeuTyThrLysLeuGlyLys 1861

602 CAGAGACCTGACAGATCGCCAGAACATGAATTTGATTTACCAAACTGGGAG 661

1862 ValProProAlaValIleIleProProAlaAlaProLeuSerGlyArgArgArgPro 1881

662 GTGCCCCCTGCTGATTATTATTCCTCCAGCTCTCTCCCTTACGAGAGAGCGAGACC 721

1882 ThrLysSerLysGlySerLysSerSerArgSerSerSerLeuGlyAsnLysSerProGln 1901

722 ACTAAAGCAAAAGGACGCAATCTAGTGAAGCAATCTCTGGGGAATTAAGAGCCCCAG 781

1902 LeuSerGlyAsnLeuSerGlyGlnSerAlaAlaSerValLeuHis 1916

782 CTTTCAGGTAC-CTGCTGTGTCAAAGTCAACTTCACTTGCAC 825

RESULT 14

AU131632 829 bp mRNA linear EST 01-AUG-2002

LOCUS AU131632 NT2RP3 Homo sapiens cDNA NT2RP3002943 5', mRNA

DEFINITION sequence.

ACCESSION AU131632

VERSION AU131632.1 GI:10991986

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo. 1 (bases 1 to 829) Ota,T., Nishikawa,T., Suzuki,Y., Ishii,S., Saio,K., Kawai,Y., Yamamoto,D., Wakamatsu,A., Nakamura,Y., Nagai,T., Sugano,S. and Isogai,T.

AUTHORS HRI human cDNA project

TITLE Unpublished (2000)

JOURNAL Contact: Takao Isogai

COMMENT Genomics Laboratory Helix Research Institute 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan Tel: 81-438-52-3975 Fax: 81-438-52-3986 Email: genomics@hri.co.jp HRI human cDNA project; 5'- & 3'-end one pass sequencing; Helix Research Institute; cDNA library construction; Department of Virology, Institute of Medical Science, University of Tokyo, and Helix Research Institute.

FEATURES

source

1..829

Location/Qualifiers

1..829

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="NT2RP3002943"

/cell_type="teratocarcinoma"

/cell_line="NT2"

/clone_id="NT2RP3"

/note="Vector: pME18SFL3; mRNA from NT2 neuronal precursor cells after 2-weeks retinoic acid (RA) induction"

ORIGIN

Alignment Scores:

Pred. No.:	Score:	Best Local Similarity:	Query Match:	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
8	65e-58	1328.00	97.46%	829	268	1	6	2	0

US-10-010-720-14 (1-2136) x AU131632 (1-829)

1737 AspLeuGluGlnAlaValIleProLysLysGluLysProGluLeuSer 1756

1 GATTGGAAACAGCTGTTCTTCTGCTGATACCAAGAGAGAGAGCTGAACCTCA 60

1757 GluProSerHisLeuAsnGlyProSerSerAspProGluAlaAlaPheLeuSerArgAsp 1776

61 GAGCTTCACATCTAAATGGCGCTCTTGACCCGAGCGCTTTTAAAGAGGAT 120

1777 ValAspAspGlySerGlySerProHisSerProHisGlnLeuSerSerLysSerLeuPro 1796

121 GTGATGATGATTCGCGTAGTCCACACTCGCCCATAGCTGAGCTCAAGAGCTTCT 180

1797 SerGlnAsnLeuSerGlnSerLeuSerAsnSerPheAsnSerSerTyMetSerSerAsp 1816

181 AGCCGAATCTAAATCAAGGCTTAATTAATTAATTAATTAATTAATTAATTAATTA 240

1817 AsnGluSerAspIleGluAspGluAspLeuLysLeuGluLeuArgArgArgAspLys 1836

241 AATAGTCAAGTATCGAAGAGAGAGAGCTTAAAGTTAGCTGCAGACACTACAGATCAA 300

1837 HisLeuLysGluIleGlnAspLeuGlnSerArgGlnLysHisGluIleGlnSerLeuTy 1856

301 CATCTCAAAAGAGATTCAGACCTCAGAGTCCAGAGCAAGCAATGAATGAAATCTTGTAT 360

1857 ThrLysLeuGlyLysValProProAlaValIleIleProProAlaAlaProLeuSerGly 1876

361 ACCAAACTGGGCGAGGTGCCCCCTGCTGTTATTCTTCCCAAGCTGCTCCCTTTCAGGG 420

Mon Sep 27 08:49:58 2004

us-10-010-720-14.rst

Page 19

Search completed: September 25, 2004, 10:29:55
Job time : 7223 secs

Blank Sheet

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 24, 2004, 00:52:38 ; Search time 37 seconds
(without alignments)

3005.996 Million cell updates/sec

Title: US-10-010-720-14

Sequence: 1 MSGGAABKQSGTSPGSLFLSP.....NISNLSKISINPPGSLNRTT 2136
10812

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	761.5	7.0	613	PKPA_PHYBL	Q01577 phycomyc
2	494.5	4.6	5703	MUSB_HUMAN	Q9HC84 homo sapien
3	481	4.4	1802	HKRI_YEAST	P41809 saccharomyc
4	475.5	4.4	2090	N214_HUMAN	P35658 homo sapien
5	460	4.3	1367	AMVH_YEAST	P08640 saccharomyc
6	448	4.1	5147	PCLO_HUMAN	Q9Y670 homo sapien
7	439	4.1	2468	MAPB_HUMAN	P46821 homo sapien
8	431.5	4.0	792	NEKA_MOUSE	Q92132 mus musculu
9	429.5	4.0	3644	MINT_MOUSE	Q62504 mus musculu
10	423	3.9	5085	PCLO_RAT	Q9JKE6 rattus norv
11	416	3.8	1080	NKRI_YEAST	Q9J119 m nucleat r
12	415.5	3.8	2067	PCLO_MOUSE	Q9SPU3 gallus gall
13	413.5	3.8	5120	MAK4_HUMAN	Q95819 homo sapien
14	411	3.8	1239	PCLO_MOUSE	Q9GYX7 mus musculu
15	404.5	3.7	5038	MAK4_MOUSE	P97820 mus musculu
16	398	3.7	1233	MAK4_MOUSE	Q02811 homo sapien
17	396	3.7	5179	MUC2_HUMAN	Q9R818 homo sapien
18	394.5	3.6	1360	TNIK_HUMAN	Q9R818 homo sapien
19	391	3.6	5560	SPEN_DROME	Q9R818 drosophila
20	388.5	3.6	865	CPN_DROME	Q9R818 drosophila
21	387.5	3.6	1306	MSB2_YEAST	P31334 saccharomyc
22	385.5	3.6	2464	MAPB_MOUSE	P14873 mus musculu
23	385	3.6	841	NEKA_HUMAN	P51957 homo sapien
24	384.5	3.6	3664	MINT_HUMAN	Q9R818 homo sapien
25	383	3.5	2453	NCR1_MOUSE	Q60974 mus musculu
26	383	3.5	2459	MAPB_RAT	P15205 rattus norv
27	382.5	3.5	1609	FIG2_YEAST	P25653 saccharomyc
28	378	3.5	939	STG2_YEAST	Q03497 saccharomyc
29	376.5	3.5	2738	PGCV_RAT	Q9R818 rattus norv
30	375.5	3.5	1308	MAK6_MOUSE	Q9JME2 mus musculu
31	370.5	3.4	2063	NCO6_HUMAN	Q14686 h nucleat r
32	365	3.4	1140	YM96_YEAST	Q04893 saccharomyc
33	363.5	3.4	4911	MLL3_HUMAN	Q8NEZ4 homo sapien

34	363	3.4	1332	1	MAK6_HUMAN	Q8N4C8 homo sapien
35	360	3.3	3358	1	PGCV_MOUSE	Q62059 mus musculu
36	358.5	3.3	794	1	KI11_HUMAN	Q8T0C3 homo sapien
37	357.5	3.3	1051	1	UKL1_MOUSE	Q70405 mus musculu
38	357	3.3	1196	1	HIK2_MOUSE	Q9QZT5 mus musculu
39	357	3.3	1223	1	KI25_CAEEL	Q10925 caenorhabd
40	356	3.3	1411	1	TCOF_HUMAN	Q13428 homo sapien
41	355	3.3	1198	1	HIK2_HUMAN	Q9H2X6 homo sapien
42	355	3.3	4903	1	MLX3_MOUSE	Q8B7H4 mus musculu
43	354.5	3.3	3866	1	HRX_MOUSE	P55200 mus musculu
44	353.5	3.3	1258	1	NEK1_HUMAN	Q96PY6 homo sapien
45	352.5	3.3	779	1	NIM1_NEUCR	P48479 neurospora

ALIGNMENTS

RESULT 1
ID PKPA_PHYBL STANDARD; PRT; 613 AA.
AC Q01577;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Serine/threonine protein kinase PKPA (EC 2.7.1.1-).
GN PKPA.
OS Phycomycetes blakesleeanus.
OC Eukaryota; Fungi; Zygomycetes; Mucorales; Mucoraceae;
OC Phycomycetes.
OX NCBI_TaxID=4837;
RX MEDLINE=96120859; Pubmed=8590476;
RA Ruiz-Perez V.L., Murillo F.J., Torres-Martinez S.;
RT "Pkpa, a novel Phycomycetes blakesleeanus serine/threonine protein
kinase."
RT Curr. Genet. 28:309-316(1995).
CC -1- FUNCTION: Serine/threonine protein kinase that probably
controlling nuclear proliferation.
CC -1- PARTICIPATES AS AN INTERMEDIATE IN AN INTRACELLULAR SYSTEM
CC -1- DEVELOPMENTAL STAGE: Mainly expressed during the vegetative
growth, the level decreased when the mycelium can differentiate
and form sporangiothecae.
CC -1- SIMILARITY: Belongs to the Ser/Thr family of protein kinases.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z46636; CA86606.1; -;
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR008271; Ser_thr_pkin_AS.
DR Pfam; PF00069; Pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; FALSE_NEG.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW Transferase; Serine/threonine-protein kinase; ATP-binding.
FT DOMAIN 17
FT NP_BIND 23
FT BINDING 50
FT ACT_SITE 140
FT BINDING 140
FT ACT_SITE 140
SO SEQUENCE 613 AA; 69594 MW; E988308B39101AIE CXC64;
Query Match 7.0%; Score 761.5; DB 1; Length 613;
Best local similarity 35.6%; Pred. No. 2.7e-19;
Matches 185; Conservative 90; Mismatches 174; Indels 71; Gaps 20;

QY 210 KAVGNSNDGRFLKDLIEIGRSGFKTVKGLDTEETVEVAMCELOPKLT--KSERORFKE 267
 Db 6 KVIKESNGRYSKLNITVLGKAYKVVYKAIDREMI-----NDNETNKKVTRQOEKXD 58
 QY 268 ---DAEMKKGLOHPIVAFYDSEWSTVKKKCIIVITELMTSGTLKTKPKVM-KIKV 323
 Db 59 LGHEIDILKSVHPIITFFHDAWVETE-----FVITELMTSGTLREYIRKLTPLPNIKI 114
 QY 324 LRSRCROILKGLQFHTPTPIIHNKDKDNIFITGPGSVKIGLGLATLKRSFAKSV 383
 Db 115 VKRMCROILKGLAYIHGHEPPIIHNKDKDNIFITGPGSVKIGLGLATLKRSFAKSV 173
 QY 384 IGTEPFAPEMYEEK-YDESVDVYAFGCMLEMAITSEYSECONAAQIYRRTSGVPA 442
 Db 174 IGTEPFAPEMYEEKYDEKVDIYAFGCMLEMAITSEYSECONAAQIYRRTSGVPA 233
 QY 443 SFDVVAPEVEIIEGCTRKNDKERYISKDLNNAFFOEELGVAVELAEDEGKXIAIKL 502
 Db 234 CLSRVQDELTLTNICTP-EDERMTAQELIERFLAVEP--EVVLVSKDMTKL---L 287
 QY 503 WLRIEDIKKLGKXKDNBAIEFSPDLERDVEDVAOEVESGVY-----CE----- 548
 Db 288 TLQV-----VFRG--MDKLAVFEFNADPTDADVAENIEQVYQNCYQDLITCEINRIL 341
 QY 549 -----GDHKTMAKAIKRVSLIKRKEQOLVRE---EQEKKKQ--EESSLK 590
 Db 342 RDIAKNOGPPDKGEDEKIVRRENDIRSELEKAKDLALAVERYFEAKKCELLKQNTI- 400
 QY 591 QOYVSSASQGTGKQPSASTGIFTASTASVSTQVEPE---EPKADHQLOLYQOPSI 647
 Db 401 --IAEERCKET--IFALQAKFQIPDLQPOPOPOPOPOPOPOPOLOQOYLSFQ 457
 QY 648 SVLSDGTVDGSGSVFTESRVSQOTVSYSGQHEQAHST 687
 Db 458 TTSPPGPTSDNSTNMTLS--LESLSKLCVSGDEQVEIT 496

RESULT 2

MUSB HUMAN

ID MUSB HUMAN STANDARD; PRT: 5703 AA.

AC 09HCG4; 000447; 000573; 014985; 015494; 095291; 095451; Q14881;

AC 099552; Q9JDE28;

DT 28-FEB-2003 (Rel. 41, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DE 28-FEB-2003 (Rel. 41, Last annotation update)

DE Mucin 5B precursor (Mucin 5 subtype B, tracheobronchial) (High

DE Mucin 5B precursor (Mucin 5 subtype B, tracheobronchial) (High

DE Mucin 5B precursor (Mucin 5 subtype B, tracheobronchial) (High

DE Mucin 5B precursor (Mucin 5 subtype B, tracheobronchial) (High

DE Mucin 5B precursor (Mucin 5 subtype B, tracheobronchial) (High

DE Mucin 5B precursor (Mucin 5 subtype B, tracheobronchial) (High

DE Mucin 5B precursor (Mucin 5 subtype B, tracheobronchial) (High

DE Mucin 5B precursor (Mucin 5 subtype B, tracheobronchial) (High

DE Mucin 5B precursor (Mucin 5 subtype B, tracheobronchial) (High

DE Mucin 5B precursor (Mucin 5 subtype B, tracheobronchial) (High

DE Mucin 5B precursor (Mucin 5 subtype B, tracheobronchial) (High

DE Mucin 5B precursor (Mucin 5 subtype B, tracheobronchial) (High

DE Mucin 5B precursor (Mucin 5 subtype B, tracheobronchial) (High

DE Mucin 5B precursor (Mucin 5 subtype B, tracheobronchial) (High

DE Mucin 5B precursor (Mucin 5 subtype B, tracheobronchial) (High

DE Mucin 5B precursor (Mucin 5 subtype B, tracheobronchial) (High

DE Mucin 5B precursor (Mucin 5 subtype B, tracheobronchial) (High

DE Mucin 5B precursor (Mucin 5 subtype B, tracheobronchial) (High

DE Mucin 5B precursor (Mucin 5 subtype B, tracheobronchial) (High

DE Mucin 5B precursor (Mucin 5 subtype B, tracheobronchial) (High

DE Mucin 5B precursor (Mucin 5 subtype B, tracheobronchial) (High

DE Mucin 5B precursor (Mucin 5 subtype B, tracheobronchial) (High

DE Mucin 5B precursor (Mucin 5 subtype B, tracheobronchial) (High

DE Mucin 5B precursor (Mucin 5 subtype B, tracheobronchial) (High

DE Mucin 5B precursor (Mucin 5 subtype B, tracheobronchial) (High

DE Mucin 5B precursor (Mucin 5 subtype B, tracheobronchial) (High

DE Mucin 5B precursor (Mucin 5 subtype B, tracheobronchial) (High

DE Mucin 5B precursor (Mucin 5 subtype B, tracheobronchial) (High

RN [4]
 RP SEQUENCE OF 1326-4895 FROM N.A.
 RC TISSUE=Placenta;
 RX MEDLINE=97166151; PubMed=9013550;
 RA Desseyn J.-L., Guyonnet-Dupert V., Porchet N., Aubert J.-P.,
 RT Laine A.;
 RT "Human mucin gene MUC5B, the 10.7 kb large central exon encodes
 RT various alternate subdomains resulting in a super-repeat. Structural
 RT evidence for a 11p15.5 gene family.";
 RL J. Biol. Chem. 272:3168-3178 (1997).
 RN [5]
 RP SEQUENCE OF 4057-4480 FROM N.A.
 RC TISSUE=Salivary gland;
 RX MEDLINE=97292540; PubMed=9147051;
 RA Nielsen P.A., Bennet E.P., Wandall H.H., Therkildsen M.H.,
 RT Hannibal J., Clausen H.;
 RT "Identification of a major human high molecular weight mucin
 RT (Mg1) as tracheobronchial mucin MUC5B.";
 RL Glycobiology 7:413-419 (1997).
 RN [6]
 RP SEQUENCE OF 4721-5703 FROM N.A.
 RC TISSUE=Gall bladder;
 RX MEDLINE=97293229; PubMed=9164870;
 RA Keates A.C., Nunes D.P., Afchal N.H., Troxler R.F., Offner G.D.;
 RT "Molecular cloning of a major human gall bladder mucin: complete C-
 RT terminal sequence and genomic organization of MUC5B.";
 RL Biochem. J. 324:295-303 (1997).
 RN [7]
 RP SEQUENCE OF 4809-5687 FROM N.A.
 RC TISSUE=Sublingual gland;
 RX MEDLINE=96125355; PubMed=8554565;
 RA Troxler R.F., Offner G.D., Zhang F., Iontcheva I., Oppenheim F.G.;
 RT "Molecular cloning of a novel high molecular weight mucin (Mg1)
 RT from human sublingual gland.";
 RL Biochem. Biophys. Res. Commun. 217:1112-1119 (1995).
 RN [8]
 RP SEQUENCE OF 4859-5703 FROM N.A.
 RC TISSUE=Placenta;
 RX MEDLINE=97347489; PubMed=9201995;
 RA Desseyn J.-L., Aubert J.-P., Porchet N., Laine A.;
 RT "Genomic organization of the 3 region of the human MUC5B mucin.";
 RL J. Biol. Chem. 272:16873-16883 (1997).
 CC -1- FUNCTION: Salivary mucin that is thought to contribute to the
 CC lubricating and viscoelastic properties of whole saliva.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- TISSUE SPECIFICITY: Expressed mainly in bronchus glands and also
 CC in submaxillary glands, endocervix, gall bladder, and pancreas.
 CC -1- PTM: Highly glycosylated.
 CC -1- SIMILARITY: Contains 1 TIL (Trypsin inhibitory-like) domain.
 CC -1- SIMILARITY: Contains 3 VWFC domains.
 CC -1- SIMILARITY: Contains 4 VWFC domains.
 CC -1- SIMILARITY: Contains 1 C-terminal cysteine knot-like (CTCK) domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 DR EMBL, AF107899; AAC36673.1; -
 DR EMBL, AF086604; AAC67545.1; -
 DR EMBL, AJ004862; CAA06167.1; -
 DR EMBL, Z72496; CAA96577.1; -
 DR EMBL, X74955; CAA52910.1; -
 DR EMBL, U63836; AAB61398.1; -
 DR EMBL, U78554; AAC51344.1; -
 DR EMBL, U78552; AAC51344.1; JOINED.
 DR EMBL, U78553; AAC51344.1; JOINED.
 DR EMBL, U78551; AAC51343.1; -
 DR EMBL, U95031; AAB65151.1; -
 DR EMBL, Y09788; CAA70926.1; -

Query Match	Best Local Similarity	4.6%; Score 494.5; DB 1; Length 5703;	Matches 385; Conservative 210; Mismatches 642; Indels 659; Gaps 82;
FT	CONFLICT	34	34
FT	CONFLICT	95	100
FT	CONFLICT	104	104
FT	CONFLICT	142	142
FT	CONFLICT	225	225
FT	CONFLICT	330	331
FT	CONFLICT	337	337
FT	CONFLICT	356	356
FT	CONFLICT	362	362
FT	CONFLICT	369	369
FT	CONFLICT	374	374
FT	CONFLICT	393	394
FT	CONFLICT	468	469
FT	CONFLICT	512	512
FT	CONFLICT	585	587
FT	CONFLICT	601	601
FT	CONFLICT	628	629
FT	CONFLICT	633	633
FT	CONFLICT	676	676
Query Match	Best Local Similarity	4.6%; Score 494.5; DB 1; Length 5703;	
Matches 385; Conservative 210; Mismatches 642; Indels 659; Gaps 82;			
596	SSASOTGICOLPSASTG----	IPRSTASASTOVEPEEPADQHOLOYOQSPISIVL	650
2502	TTPTATSVTPRIPSSSLGTTWRLSQTTTPATATSTAPSSSTPETAH-----	ISTVL	2552
651	SDGTVDGSGGSSVFTESS-----	RVSSQOIVSYSGSHEQAH-----	SNGTIV 691
2553	TATATTCATGSVATPSSTPGTAAHTTAVPTTTTGGFATPSSSGTALPWPWISTTTP		2612
652	GHIPSTVQAQGRHGV-----	PRSSVAG-----	OSQGPS--SALTGVSS 733
2613	TTRGSTVPSISPGTTHNAVLTITTTTVAATGSMATSSSTQISGTPSLTITATITAT		2672
724	QPIHQPOQOQIQAOPQOIVQYLSQTSSEATTAQOPVOPADQVLPVVSAGKOST		793
2673	GSTINPSSPTGTRIPPLVLT-----	ATTPAATSVITPSSALGTTHTPPV--PNTA-----	TT 2726
794	QGVSAQVAPAEPAV-----	AQPAOTPTTLASSVDA-----	HSVYASG-- 832
2727	HGRS--LSPSSPHVTATMTATSGTLGTTHTBESGTSHTPATTTGTHSHPPALSSPH		2785
833	-MSDGENVPSSSGHGEKT--	TRHVKXRSR-----	HEKTSRPLRLINVS-- 880
2786	PSRRTESPSSPGTTPHTTATSKTTATAPSKRTISTLLPSSPTAPITTVTMGCEP		2845
881	-----	NKGD-----	RV 886
2846	QCAWSHMLDYSPMGPSGDPEDTYSNIRAAGAVCEQPLGLECRQAQOPVPLRELGOV		2905
887	VEQOLE-----	THNRKV-----	TEKFDD--
2906	VECSLDPLGVLVJCNRRQVQKFKMCFYVEIRVFCNNGHCSPAT-----		2949
932	VDOVRRIIEKADMLSEDVSEVPEDDQ-----	LESLOKQDYGSGSOKLEGEFKOPIP	986
2950	-----	SSTATPSSPTGTTWLTLEQTAAITTAATTAATGSTALPSSPTGPAP	2992
987	ASS--MPQOIGIPTSLQVVSAGRRFV-----	SPVESRL-----	1022
2993	PPKVLISQATTPPATSSKATSSSPRLATLVLNSTAKTATSFPIPSSTIGTGT		3052
1023	-----	RESKVPSEITDVAASSTAOSSPGKMLSHSASSLSLOA-----	FSELR 1066
3053	QNRPHPMAMSTIHSSPETHSTVLTATATTTATSSSTSSPTGTTWLTLELT		3112
1067	AQMTGEG--PNTAPNNSHTGPFPPVPLPSSLAGPTTPAATAVAPATSSPNDISTSV		1124
3113	AATTTAALPHGTS--STPGTTWLTLEPSTTAVTVPTGSTATAS-----		3155

FT CARBOHYD 1400 1400 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 1802 AA; 188890 MW; E344CA6469785A24 CRC64;
 Query Match 4.4%; Score 481; DB 1; Length 1802;
 Best Local Similarity 22.6%; Pred. No. 2.3e-09;
 Matches 333; Conservative 210; Mismatches 603; Indels 328; Gaps 64;

QY 596 SGAAGTGKQPLPASTGIPITASTASVSYQVEPEPEADHQOL--QYQGSISVLSD- 652
 DB 29 STSYNNGIESTPSYST--SAISSTGSNKENAITSSETTTMAGQYGEAGSTTIDE 83
 QY 653 -GTVDGSGSSVFTESRVSSQOTVSYGSGHEQASHGTVPCHI PSTVQAQSGPHGV---- 707
 DB 84 QETGHSQYISVTTTGTQTDIMS-SYKSTELIATPESSI--VPTLQGSYDESQISQTL 139
 QY 708 -YPPSSVQAQSGQGPS-SSSLTVGASSQPIHQPOQOQGIQOTAPQOQVQVLSQTS 765
 DB 140 SHNPKSVASDSDTSTSESSSSVILSTSDSAVPREISPIITTDQISKEEGTLAQTSI 199
 QY 766 SEATTQAQVQQAQVLPQVAGKQSTQGVQVAPAEVAVAQPOATPTLASV-- 822
 DB 200 SEIT-----RIAGVTRVSO-----ISITASTIDGSSSTO-TDPSNTVSE 243
 QY 823 DSAHSDVAGMSDGNENVPSSSGRHEGRTTKRHYKSVSRSEKETSBPRLIILVSNK 882
 DB 244 NSVEEYAMSKQLSSTSYSSSTVYSG-----GESTADKTSSTP--ITSFSS 289
 QY 883 GDRVECOLETHNRKRVTEKFDLDGNPEIATIMVNDPILAESEFPDQREIIEKA 942
 DB 290 YSQTTSTETSESRVAV-----GVSRPSTTQTTSI-- 320
 QY 943 DEMLSDEVYEPBGDGLSLQKDYGFSGSQKLEGEFKQPIPASMPQOI-----GI 996
 DB 321 DSFMSSEVELSTY-----YDLSAGNYPDOELIVDRPAISSTIETSESAQGV 367
 QY 997 PTSLSLTQVHVS-AGRRITVPVPEBRLREKVPSEITDVAASTAQSPGMNLHSAASL 1055
 DB 368 SRESNTAVASISITNTIVSSASDT-----VISTSTNTVPSSVMTVHAITSSTYL 421
 QY 1056 SLOQAFSELRRQWTEGPNTAPPNFSHTGTFPVVPPFLS--SIAGVPTTAATAVPA 1112
 DB 422 S-----SSLYSSPFLSASVSHFG-----VAFPPAYISFSSVPAVAST--Y 462
 QY 1113 TSSPPNDISTSVIOSEVTPPEBGIAGVANTGVVTSGGGPIPVSSPVLSSVSSITI 1172
 DB 463 TSSP-----SASVVP-----SAYAS-----PSVPA--VSTTSSPSA 496
 QY 1173 PAVSIS-TTSPSLQVPTSTSEIVSS--TALYPSVTVSNTSASAGSTATPGKPPA 1227
 DB 497 PAIISSTYSSPS--APVAVSTYSSPSAPAIISSTYSSPSAPVAVSSTYSSPSAPA 554
 QY 1228 VVSOQAAAGSTTVGATLTSVSTTSPPTASQSLQSSSTPTPLAETVTVSASLDKTS 1287
 DB 555 AISTYTSSTPS--APVAVSTYSSPSAPAIISSTYSSPSAPVAVSSTYSSPSAPAI 612
 QY 1288 HSGTTG-----LASLSAPSSSSSPGAGVSYIQPGGLHVLPSVIASTPILPQAAGP 1342
 DB 613 SSTYSSPSAPVAVSSTYSSPSAPAIISSTYSSPS--VPAVAVSSTYSSPSAPAISS 670
 QY 1343 TSTPLPQVPSIPPLVQV--ANVPAVQOTLIHSQOPALLPNOPTHQPEVDS-- 1394
 DB 671 TYT-----SSPSVPAVAVSSTYSSPSAPAIISSTYSSPSAPVAVSSTYSSPSAPA 726
 QY 1395 --DTQPKAPGIDIKTLEKRLSFLSEHSSGCAQHASVLESLVYESTVPGIPTTVA 1452
 DB 727 TYTSSPSAP-----VAVSSTYSSPSAPAIIS--STYSS--PSAPVA 765
 QY 1453 PSKLTISTTSTCL--PTTNPLGTVALPVT-PVTVGQVSTPVSTTSGVKGPGTAPSK 1507
 DB 766 VSTYSSPSALVAVSSTYSSPSVYIVSPSTFAAISGYTSP--SASVAMSTSSSS 822
 QY 1508 PLTKAPVLPVGTETLPAGLTPEQQLPFGPSLTOSQOPLDLDAQLARTTSPEXITVTS 1567

DB 823 P--YDIYVSLSSASRSSTATYEFSPSPSTSLPTS-STYTFYSSAVAFESPESERYSTTS 878
 QY 1568 AVGPVSM-AAPTATTEACTOPQKGVSYQKEGPVIAL-----SSGAGVFEMKGFQVSV 1618
 DB 879 TIAPQIHSTLSRLITDFLQTSMAIQSVSQOISHTSTLNDIEHSSALSVF----- 929
 QY 1619 AADQAKGKKKSKEDAKSVHEESSISESSVSSSSPESTLVKPEPNGITIPGISDVPS 1678
 DB 930 -----NPS-----ASNIVETSLIIST-QASITSPK-NSAKISLQSQLSSS 969
 QY 1679 AHK--TTASEAKSDPQTKVGRFOVTTTANKVGRFVSYSKTEKTDTRKKEGPVAPPPM 1736
 DB 970 TKNPYDTANKNTEISGRSTVSNFYTSSAKPKDNKESATPTELT----- 1015
 QY 1737 DLEQAVLPAVLPKKEKPELSEPS--LNGPSDPEAFLSKVDVDDGSGSP--HSPHQ 1789
 DB 1016 -----TSSSHAYSLIFSHSNVYGLSHNPFVDSKSNATSGFSSSS 1058
 QY 1790 LSKSKLPSQNL--SQSLN-----SFNSYWSNDNEIDEDLKLRLRDKHLKEI 1841
 DB 1059 ISSIKLSKETIPASKSVSNTQERITSFTSLRAN--SQSEKSEGNVSGSLQSHISSN 1115
 QY 1842 QDLSQKHEIRISVTKLGKVPFAVILPPAPLGRRRPRTSKSGSSRSSSLGKNSPO 1901
 DB 1116 PELSTTKVDKSKSLSKKSKYKMGF-----NGETGLTTTKQYKSSSTSGSYSS 1166
 QY 1902 LSGNLGSGASAVLHPQTLHPGNIPESQNLQLOPKPSPSDNLVYAFITSDGALSP 1961
 DB 1167 FT-KISIGPATVAVQASTNSVETAPALS--TYPTTPYP-SPNSVAMLPV--AIIVE 1218
 QY 1962 SUSAPGQGSTNTVATVANSQAQAPAMTSS 1995
 DB 1219 S-SETGPTTASFN--PSITGLPNAIEPAVAVS 1248

RESULT 4
 N214 HUMAN
 ID N214 HUMAN STANDARD; PFT; 2090 AA.
 AC P35658;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Nuclear pore complex protein Nup214 (Nucleoporin Nup214) (214 kDa
 NUP214 OR CAN OR CAIN.
 GN Homo sapiens (Human).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Testis;
 RX MEDLINE=92195315; PubMed=1549122;
 RA Von Lindern M., Fornerod M., Van Baal S., Jaegle M., De Wit T.,
 RA Bulj A., Grosfeld G.,
 RT "The translocation (6;9), associated with a specific subtype of acute
 RT myeloid leukemia, results in the fusion of two genes, deK and can,"
 RT and the expression of a chimeric, leukemia-specific deK-can mRNA".
 RL Mol. Cell. Biol. 12:1687-1697(1992).
 RN [2]
 RP CHARACTERIZATION.
 RX MEDLINE=94151361; PubMed=8108440;
 RA Kremer D., Wozniak R.W., Blobel G., Radu A.;
 RT "The human CAN protein, a putative oncogene product associated with
 RT myeloid leukemogenesis, is a nuclear pore complex protein that faces
 RT the cytoplasm.".
 RL Proc. Natl. Acad. Sci. U.S.A. 91:1519-1523(1994).
 CC - FUNCTION: May serve as a docking site in the receptor-mediated
 CC import of substrates across the nuclear pore complex.
 CC - SUBUNIT: Homodimer. Interacts with DDX19 and NUP88.
 CC - SUBCELLULAR LOCATION: Nuclear pore complex. Cytoplasmic filaments.
 CC - TISSUE SPECIFICITY: Expressed in thymus, spleen, bone marrow,
 CC kidney, brain and testis, but hardly in all other tissues or in

CC whole embryos during development.
 CC -1- DOMAIN: Contains F-G repeats.
 CC -1- PTM: Probably glycosylated as it reacts with wheat germ agglutinin (WGA).
 CC -1- DISEASE: Implicated in a subset of acute myeloid leukemia (acute nonlymphocytic leukemia) (AML) carrying a chromosomal translocation t(6;9) (p23;q34) that results in the formation of a DER-CAN fusion gene.
 CC -1- DISEASE: Involved in some cases of acute undifferentiated leukemia (AUL) through a chromosomal translocation t(6;9) (q21;q34.1) that involves NUP214/CAN and SET.
 CC -1- DATABASE: NAME=Atlas Genet. Cytogenet. Oncol. Haematol.;
 CC -1- WWW="http://www.infobiogen.fr/services/chromancer/Genes/CAN.html".
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or_send_email_to_license@isb-sib.ch).
 CC -----
 CC EMBL: X64228; CAA45535.1; -
 CC PIR: S26058; S26058.
 CC Genew: HGNC:8064; NUP214.
 CC MIM: 114350; -
 CC GO: GO:0005643; C:nuclear pore; TAS.
 CC GO: GO:0005215; F:transporter activity; TAS.
 CC InterPro: IPR004325; Nucleoporin_Fg.
 CC InterPro: IPR01680; WD40.
 CC Pfam: PF03093; Nucleoporin_Fg_22.
 CC SMART: SMO0320; WD40; 2.
 CC KW Nuclear protein; Transport; Proto-oncogene; Chromosomal translocation; Repeat; Glycoprotein.
 CC FT DOMAIN 481 2076 11 X 5 AA APPROXIMATE REPEATS.
 CC FT DOMAIN 1409 2084 18 X 4 AA APPROXIMATE REPEATS.
 CC FT DOMAIN 1427 2085 11 X 3 AA APPROXIMATE REPEATS.
 CC FT DOMAIN 1213 2090 PRO/SER/THR-RICH.
 CC FT DOMAIN 740 768 LEUCINE-ZIPPER 1.
 CC FT DOMAIN 861 882 LEUCINE-ZIPPER 2.
 CC SITE 812 813 BREAKPOINT.
 CC SEQUNCE 2090 AA; 213766 MW; 6DBE767FDD857F8F CRC64;
 QY Query Match 4.4%; Score 475.5; DB 1; Length 2090;
 Db Best Local Similarity 20.7%; Pred. No. 4.2e-09;
 Matches 392; Conservative 239; Mismatches 740; Indels 525; Gaps 79;
 QY 545 YVCEGPHKTKAKIXORVSLIKKKRQ-----RQVREDEQKKKQESSIKQVEGS 596
 Db 309 YIEEMVLVLAASASTEVSTILAKOSQIWNESWLBEDSSRAELPVYDKSDSDSLPMGVVD 368
 QY 597 SASQTGI-----KQLPSA-----STGIPTASTASVSTOYE--- 628
 Db 369 YINQVETITSEKTLPPAPVLMSTDVLCPPYMINQNGVKSILIKTERISLSEEROP 428
 QY 629 -----PEEPBADQOC-----LQYQPSISVLSDETVDGQSSSVFT--ESR 668
 Db 429 KSPGSTPTTTSQAPQKLDASAAAPASLPPSSPAPIATFTSLIPAGAPVTFSGSSS 488
 QY 669 VSSQQTV-----SYSGNHEQASHSTGVPHISTVQAGQPHGVPPSSVA---QGSGQ 719
 Db 489 LKSSATVTGEPSPSSGSD--SSKAAPRPGPST-----FSFVPPSASALPAPPAAP 538
 QY 720 GQPSSSSILTVSS-----SQPIHPQOCQ--GIQOTAPQ--QTVQYSLSCSTSSSEATT 770
 Db 539 VAPSAASFSGSSGFKPTESTPVPVSAPNAPAMSSFFPSPSAVKVNLSEKFTL--AAT 596
 QY 771 AQPVQPC--APVYLQVSAKQSTOG--VSQVAP--AEPVAAQOPATOPTTLASSVDSAH 826
 Db 597 STPVSSSSAPRPMSPFSSASKAPLSPHPTPLAPPSVPLKSSVLPSPSGRAGSS 656
 QY 827 SDVASGMDGEMENVSSSGRHGRTTKKHVRSVARSRHE--KTSRPKRLILNVSNKDR 885

Db 657 SPVSEVMVCKSPRIITPPA--KFGSPQAKSLQPAVAEKQHQWCKSDPVM-----AGIGHE 709
 QY 886 VVECOLETHNRKRMVTFEKEDLDGDNPEEIAITIVNND-----FYIAER----- 928
 Db 710 IAHQKELEELKARISKACQFQVIGSEKKMLRTBEDDLITLEIKETTESLHGDISSLK 769
 QY 929 -----ESF--VQVREIITKADQM-----LSEVSVPEPEDQGLSEIQKDYDFSGSQK-- 976
 Db 770 TTLLEGFAGVEAEARQERNRNDSGYLHLYRMRPLDPKSEAQIOETRRHQYVKFAVQVNV 829
 QY 977 --LEGFEKQPIPPASSMPQOIGIP-----TSLTQVNHAGRRF--TVSPVPSRLRE 1024
 Db 830 DVLDLMDQHLQKQKQKQHLVPERETLFTNLANNREIINQKRLNHLVLSLOQLRYK 889
 QY 1025 -----SKVPSSE-----ITDVAASTAGSPGNLSHSASLSLQQAFLSEL 1065
 Db 890 QTSLSMSLSAVPSQSSIHSPDSELSCLNALKTTIE-----SHTSILPRVPAKLSFMK 943
 QY 1066 RAQM-----TEGP--NTAPNFSHTGPTTPVVPPLS-----SIAGVPTTAAT--- 1107
 Db 944 QOQLRNLFLARKTPPVBSTAPASLSRSA-----FLSQRYEDLDEVSTSSVSGGLE 995
 QY 1108 -----APVPATSSPNDISTVY-----QBEVTVPTRETAG-- 1139
 Db 996 SEDARTCKDDEAVVQAPRRHAPVYRTPS-----IOPSLPHAAPAKSHLVHGSSPGVMGT 1051
 QY 1140 -VATSTG-VVTSG-----GLPIP--PV-----SESPVLSV 1166
 Db 1052 SVATISAKKIIPQAGADSTMTLTKTVKGAAPSPHSHISAPQQLAAALRRQMAQAPAVNTL 1111
 QY 1167 VSSI--TIPAVVISITTSBLSQVP--TSTSEIVVSTLAYS---VTVASATSAGSGTAT 1220
 Db 1112 TBSLTKNVPQVAVVQELKNNPATPSTAMGSSVYSTAKTGHVPLTPVAANQAKQGLHNS 1171
 QY 1221 PGKPPPAVVVSOQAAGSTVGCATLTSVTTSPSTASQSLQSSSTSTTLAETVVA 1280
 Db 1172 LKSGGTPPASGQSSQSDKASGTAKIETAVTSTBSAGQSKPSPFSFGTFNGIITPT 1231
 QY 1281 HSLDKTSHSSTTGLAFSLASPSSSSPGAGVSSYISQPGIHLVLPVSLVSPILPOAA 1340
 Db 1232 PSENFTMAOGAVPTSTKSSQPDAFSSGGGSKPEYEAIPSSPSPGSLTASNTTPRGEAAS 1291
 QY 1341 GPTSPPLLPQVPSIPPLVQPVANVPA--VOOTLIHSGOPALPENOHTHCEVDSPTOP- 1398
 Db 1292 --SSRVAASGALSTSSKLETPPKIGELLPSSLAGETLGSFGLRGVQADSDTKPT 1349
 QY 1399 -KAPGIDDIKLEEKLSLSEHSSSG-----AQHASVLSLTSVITSTVPGIPTTAAVP 1453
 Db 1350 NKASSTSLNSTQPTKTSQVPSGPNFTAPVVLGKHT-----BPVTSATTTSVAP 1399
 QY 1454 SKLLSTSTSTCLPPTNLPIGTVALPVPVPTPGQVS--TPVST-----TTSQVY 1500
 Db 1400 -PAATSTSTAV-----FG--SLPVTASGSGVLSFGGTSLSAGKTSFSGSQQINSTV 1450
 QY 1501 PGTAAPSKPLTKA-----PVLVGTLEPAGTLSEQLPPEPPGS--LTSQQLLEDL 1550
 Db 1451 PPSAP--PPTTATPLPUSFTLSFGSLSSATPTS--LPMAGHSTEBATSSALPEKPG 1506
 QY 1551 DAQLRRLTSPEXITVTSVAVGVPVMAAPTAITEAGT-----OQKGVQYKBEVGLATSSG 1605
 Db 1507 DSE-----VSASAASILEEQGSAQLPQAPPTQTSYDSYKKEPVL----- 1543
 QY 1606 AGVFMMKGFQVSVAADGQKBEKKNSEDAKSVHPESTSESSVSLSSSPES--TLVKPEPN 1664
 Db 1544 -----AQPAVSNAGTAASSTSLVALAABEATPATVGPDPARTEAVPPAS 1586
 QY 1665 GITTPG-----ISSVPESAKHTTASAKSDP-----GQPTVGFQVTTANKV-- 1709
 Db 1587 SFSVPGQVTAALASISAGPAVETSTSTPILASTTSIYAPGPSAEMAAFGTVTSSGSYFA 1646
 QY 1710 --GFSVSKTEDKITDTKKEGFPVASPPFMDLEQAVLPVAVIKK-----EKPELSEP 1758
 Db 1647 QPAASSSSANQLTNNTATATAPSATPVFGVAASTAPSLFGQQTGTSASTAAATPQVSS- 1705


```

QY 1759 SHLNGPSSDPEAFSLRDVDSGSPHSHQLSKSLPSQNLQSLNSFNSSYMSDNE 1818
DB 1706 SFPSSBPAFTTAPVPGGOTTFOQASV-----FCQASSASAVFSFSGQGF 1750
QY 1819 SDIEDEDLRLRLRDLKELKEIDLOSHQKHEISLTKLGVPAVILPPAALSGRR 1878
DB 1751 SSV-----PAFQOPASS----- 1762
QY 1879 RRPYSKGS-----KSSRSSLSGKNSPQLSGNLSGQASAVLHPQTLHPGNIPES 1930
DB 1763 -PPTSTGSGVFGAASSTSSSSSFSGQSSPNTGGILFGQS-----NAPAF 1806
QY 1931 GGNOLLQPLKPPSSSNLYSAFTSDGALVPSLS-----APGQSTSTNTY----- 1976
DB 1807 GQ-----SPFGQSGSVFGGSGVFGGSGNTGRGGGFGSGLGKXPSQDAANKNPSSASGFG----- 1910
QY 1977 -GATVNSQAQAQAPPMATSSRKGTFTDHLKLVDMNARDAMLGRRGSKGHNNYEGPGM 2035
DB 1858 GGVFGQSSSSSSGSGVFGGSGNTGRGGGFGSGLGKXPSQDAANKNPSSASGFG----- 1910
QY 2036 ARKFSAPGQLCTSMSTNLGSGAPISAAATSLGHFTKSMCPPOVGFPAITPFGAQMSTG 2095
DB 1911 -----GSTATSNSTNLFQGN-----SGAKTFGGFASSSFGEGK-----PTGTFSGGG 1952
QY 2096 GPAPPLGQFQFVGTASLQNFNISNLQKXISNPPG 2130
DB 1953 SVASQGFSSSPKKTGGFGAAPVPGSPPTGSSPG 1987

RESULT 5
AMTH YEAST STANDARD; PRT; 1367 AA.
ID AMTH YEAST STANDARD; PRT; 1367 AA.
AC P08640; P08068;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Glucoamylase SL/S2 precursor (EC 3.2.1.3) (Glucan 1,4-alpha-
GN glucosidase) (1,4-alpha-D-glucan glucohydrolase).
DE STAI OR STAZ OR MAUS OR YIK019C.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C / AB972;
RX MEDLINE=97313266; PubMed=9169870;
RA Churcher C.M., Bowman S., Baddock K., Bankier A., Brown D.,
RA Chillingworth T., Connor R., Devlin K., Gentles S., Hamlin N.,
RA Harris D.E., Horenell T., Hunt S., Jagsels K., Jones M., Lye G.,
RA Moule S., Odell C., Pearson D., Rajandream M.A., Rice P., Rowley N.,
RA Skelton J., Smith V., Walsh S., Whitehead S., Barrett B.G.;
RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome IX.";
RL Nature 387:84-87(1997).
RN [2]
RP SEQUENCE OF 1-242 AND 762-1331 FROM N.A.
RX MEDLINE=87194600; PubMed=3106330;
RA Yamashita I., Nakamura M., Fukui S.;
RT "Gene fusion is a possible mechanism underlying the evolution of
RT STAI.";
RL J. Bacteriol. 169:2142-2149(1987).
RN [3]
RP SEQUENCE OF 1-31 FROM N.A.
RC STRAIN=SPX101-1C;
RX MEDLINE=89031230; PubMed=3141213;
RA Pardo J.M., Ianez E., Zalacain M., Clares M.G., Jimenez A.;
RT "Similar short elements in the 5' regions of the STAZ and SGA genes
RT from Saccharomyces cerevisiae.";
RL FEBS Lett. 239:179-184(1988).
CC -1- CATALYTIC ACTIVITY: Hydrolysis of terminal 1,4-linked alpha-D-
CC glucose residues successively from non-reducing ends of the chains
CC with release of beta-D-glucose.

```

```

CC -1- SIMILARITY: TO S. POMBE SPBC215.13.
CC -1- SIMILARITY: SOME, TO S. POMBE SPCC285.13C.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL: Z38061; CAA86176.1; -
DR EMBL: M16164; AAA35014.1; -
DR EMBL: M16165; AAA35015.1; -
DR EMBL: X13857; CAA32069.1; -
DR PIR: S48478; S48478.
DR GeneOnline: 139731; -.
DR SGD: S0001458; MUC1.
DR GO: GO:0005886; C:plasma membrane, IDA.
DR GO: GO:0030447; P:filamentous growth, IDA.
DR GO: GO:0007125; P:invasive growth, IMP.
DR GO: GO:0007124; P:pseudohyphal growth, IMP.
KW Hydrolase; Glycosidase; Polysaccharide degradation; Glycoprotein;
KW Signal; Multigene family.
FT SIGNAL 1 21
FT CHAIN 22 1367
FT DOMAIN 210 1367 SER/THR-RICH.
FT CARBOHYD 817 817 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 874 874 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 1367 AA; 136110 MW; 91C00E2BDD61AA9D CRC64;

Query Match 4.3%; Score 460; DB 1; Length 1367;
Best Local Similarity 22.4%; Pred. No. 8; 8e-09;
Matches 302; Conservative 164; Mismatches 540; Indels 342; Gaps 54;

QY 585 EESLSKQVHQSS--ASQCIKQLPSASGICPASTTSS-----VSTQVPEPEAD 635
DB 229 ESSSTTSSTSSSTSSSTSSSTSSSTTAPATPTTSTCKEPTPTTSTCKEPTTP 288
QY 636 QHQQLQVQPSISVLSDGTVDGSGGSSVFTESSVQQTIVSYGSHQHAHSTGTVPCHP 695
DB 289 HHD-----TTPCKTKTTSTKCT-----KXTTPVPPTSS 319
QY 696 STVQAQSGPHGVYPSSVAQGGQGGQSSSLGVSSSQDTQHPQOQGGIQAAPPGQTV 755
DB 320 STSSSAP--VPTSSSTSSSAPVTSSTT-ESSSAPVPTSSSTSSSAP----- 370
QY 756 QYLSQTSSTSEATTAQVQAPQVLPQVSAQKOSTGVSQ--VAAPVPAVAQPOA 812
DB 371 -VTSSTSSSAPVTSSTSSSAPVPTSSSTSSSAPVTSSTSSSAPVTSSTSS 429
QY 813 TQPTTAAVSQASDVASGMSDGN--VPSGSGRHEGRTTKRHYKSKVSRSHKXT 869
DB 430 SSAPVTSSTSSSAPVTSSTSSSAPVPTSS-----TT-----ESS 469
QY 870 SREPLRIANVSNKGRVVEQLETHNRKMTFKEDLDGNDPELATIMANNDFILAIERE 929
DB 470 SAP-----VTSSTSSSAPVPT-----PSSSTT-----ESS 496
QY 930 SFVQVQVEIEIKADENLSEDEVSEPEGGQGLSLQKDDYGFSGSQKLEGFKOPPASS 989
DB 497 S-----APVTSSTSSSAPVPTP-----SSTSSSAPVPTSSS 533
QY 990 MPQOIGIPTSLTQVHSAGRRFVSPVESRLRESKVFP--SEITD-----TVAAT 1040
DB 534 TTSSSAPVTSSTSSSAPVPTSSSAPVPTSSSAPVPTSSSAPVPTSSSAPVPTSSS 588
QY 1041 AQSPEMNLISHASSLSLQAFSELRAQMTGCPNTPNFHT--GFTPPVVPPL-----S 1095
DB 589 TESSSAPVPTSSSAPVPTSSSAPVPTSSSAPVPTSSSAPVPTSSSAPVPTSSS 648
QY 1096 SIACVP-----TVAATAPVATSPPDISTSVQGSVTVPTGEGVAGTSGVTVTSG 1150

```

```

Db 649 SSAPVPTSSSTSSAPVPTSSSTSSAPVTSST---TESSAPVTSST--TESS 703
QY 1151 GLPIPVESPVLSVSSVITIPAVVSTSTSPSQVPT-----STSEIVSSALXP 1203
Db 704 SAPVPTSSSTSS---SSAPVPTSSSTSSAPVPTSSSTSSAPVTSSTSS 760
QY 1204 SVTVSATASAGGATATGPKPAVVSQAAGSTTGAATLSTSTSTSTSTSTSTST 1263
Db 761 SAPVPTSSSTSSAPVPTSSSTSSAPVPTSSSTSSAPVPTSSSTSSAPVPTSS 820
QY 1264 SSSTSTPLAETVAVVSAHSDTKSHSTTGAFSLASASSSSSSPQAGVSSYISQPG 1323
Db 821 SAPSTPTSSST---ESSVPVPTSSSTTE---SSAPVSSSTSSVAP----- 865
QY 1324 LVIPSVIATPILPQAAGPTSTPLLPQVPSIP--PLVQPVAN--VPAVQ--QTLHSQ 1379
Db 866 --VPTSSSSNITSSA--PSSIPSTSTSTSTSTSTSTSTSTSTSTSTSTSTST 921
QY 1380 LIPNQHPTCEVUSDTPQKAPGIDDKTLERLSLSESSSSGQAQHASVLETSVIE 1439
Db 922 IVPTK-----TTTSVPTSTTTTITTTVCSTGNSAGETSGCSPKTV----- 963
QY 1440 STVPGIPTTAVAPSKLSTSTSCLPPTNPLGTVAPLPTPVPTQVST--PVSTT- 1495
Db 964 TTVVPTTTTSTSTSTSTTTTITTVCSGTN---SAGETTSQSPKTTTITTVCS 1018
QY 1496 -TSGVKGTAPEKPKPLTKAPVLPVTELPAGTLPSEQLP-----PQPSLQSQPLE 1548
Db 1019 ETASESTTSTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPT 1071
QY 1549 DLDAQRLRTLSPEKITTSAVGPVMAAPAIT---EAGT---QPOKGVQVKEGPVL 1600
Db 1072 -----TTPSVTVMPTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPT 1118
QY 1601 ATSSGAVFVKGFQVQVAA-----DGAQKGRKKS--EDASVHPESTSGSSV 1648
Db 1119 PCGTGCEYTTTATLTAVTTTAVTTTSTSTSTSTSTSTSTSTSTSTSTSTST 1174
QY 1649 LSSSSPESTLVKEPENGITIPGISDVPESSAHKTASAKS----- 1689
Db 1175 LAPSAPVTPATNAVPTTITTESAATMAGETTSVCSAKTITTSASAGNTAPSAT 1234
QY 1690 DTGQPKVGRFQYTTANKGRFSVSKTDXIT-----DTKKEGPVAPSPMDLEQ 1740
Db 1235 TTAIPF-----TITTESVGTNSAGETTTGYTTKSIPTTYITLLPGSNGAKNYET 1289
QY 1741 AVLPAYIPKKEK-----PELSEPHLNGBSSDPEAALSRDVGSSSPHSPPQ 1789
Db 1290 ANTPISIKTTSQATTASASAVPVTSBS-LTGP-----LQASAGS-----A 1331
QY 1790 LSSKSLPSONLSQSLNSFNSSYSMSDN 1817
Db 1332 VATYSVP-----SISSTYQGAN 1349

```

RESULT 6

```

PCLO_HUMAN          STANDARD; PRT: 5147 AA.
ID PCLO_HUMAN          Q9Y6V0; O43373; O60305; Q9BYC8; Q9YIV2; Q9Y6I9;
AC Q9Y6V0; O43373; O60305; Q9BYC8; Q9YIV2; Q9Y6I9;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Piccolo protein (Acronin) (fragments).
OS PCLO OR ACZ OR KIAA0559.
GN Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE OF 1-759 FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=99439764; PubMed=10508862;
RA Wang X., Kibschull M., Laue M.M., Lichte B., Petrasch-Parwez E.,

```

```

RA Kilimann M.W.;
RT "Acronin, a 550-kd putative scaffolding protein of presynaptic active
RT zones, shares homology regions with rim and bassoon and binds
RT profilin."
RL J. Cell Biol. 147:151-162(1999).
[2]
RP SEQUENCE OF 552-4404 FROM N.A.
RA Kremer U., Wollam C., Woldmann P., McGrane B.;
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
[3]
RP SEQUENCE OF 3619-5147 FROM N.A. (ISOFORM 2).
RC TISSUE=Brain;
RX MEDLINE=98290545; PubMed=9628681;
RA Nagase T., Ishikawa K.-I., Miyajima N., Tanaka A., Kotani H.,
RA Nomura N., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. IX.
RT The complete sequences of 100 new cDNA clones from brain which can
RL code for large proteins in vitro."
[4]
RP SEQUENCE OF 4405-4439 FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Helel F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udell T.B., Toshyuk S., Carninci P., Prange C.,
RA Raha S.S., Loggiano N.A., Peters G.J., Abramson R.D., Mulhaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Botterfield J.S.N., Krzywinski M.I., Skalska U., Small D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
[5]
RP SEQUENCE OF 4405-5147 FROM N.A.
RA Kallio J., Elliott G.;
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: May act as a scaffolding protein involved in the
CC organization of synaptic active zones and in synaptic vesicle
CC trafficking (By similarity).
CC -!- SUBUNIT: Interacts with Rabac1, Pral, RIMS2 and profilin (By
CC similarity).
CC -!- SUBCELLULAR LOCATION: Concentrated at the presynaptic side of
CC synaptic junctions (By similarity).
CC -!- ALTERNATIVE PRODUCTS:
CC Event-Alternative splicing. Named isoforms=2;
CC Comment=Additional isoforms seem to exist;
CC Name=1;
CC IsoId=Q9Y6V0-1; Sequence=Displayed;
CC Name=2;
CC IsoId=Q9Y6V0-2; Sequence=VSP_003923, VSP_003924, VSP_003925,
CC VSP_003926, VSP_003927;
CC -!- DOMAIN: C2 domain 1 is involved in binding calcium and
CC phospholipids. Calcium binds with low affinity but with high
CC specificity and induces a large conformational change.
CC -!- SIMILARITY: Contains 2 C2 domains.
CC -!- SIMILARITY: Contains 1 PDZ/DRH domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

```

CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 DR EMBL; Y19188; CAB60727.1; -
 DR EMBL; AC004903; AAD20936.1; -
 DR EMBL; AC004886; AAD21789.1; -
 DR EMBL; AB011131; BAA25485.1; -
 DR EMBL; BC001304; AAH01304.1; -
 DR EMBL; AC004082; AAB97937.1; -
 DR PIR; T00634; T00634.
 DR HSSP; P04410; 1A25.
 DR GeneW; HGNC:13406; PCLO.
 DR MIM; 604918; -
 DR GO; GO:0005856; C:cytoskeleton; NAS.
 DR GO; GO:0045202; C:synaptic junction; ISS.
 DR GO; GO:0005509; F:calcium ion binding; ISS.
 DR GO; GO:0005544; F:calcium-dependent phospholipid binding; ISS.
 DR GO; GO:0005522; F:profilin binding; ISS.
 DR GO; GO:0007010; P:cytoskeleton organization and biogenesis; ISS.
 DR GO; GO:0016080; P:synaptic vesicle targeting; ISS.
 DR InterPro; IPR000008; C2.
 DR PRINTS; PR00360; C2DOMAIN.
 DR PRINTS; PR00399; SYNAPTOTAGM.
 DR SMART; SM00239; C2_2.
 DR PROSITE; PS00459; C2_DOMAIN_1; 1.
 DR PROSITE; PS50004; C2_DOMAIN_2; 2.
 DR Calcium/phospholipid-binding; Zinc; Metal-binding; Zinc-finger;
 KW Repeat; Alternative splicing.
 FT NON TER 1
 FT DOMAIN 400 465
 FT 10 X 10 AA TANDEM APPROXIMATE REPEATS OF
 FT P-A-K-P-O-P-O-P-X.
 FT ZN FING 499 523
 FT ZN FING 969 992
 FT NON CONS 1010 1011
 FT C4-TYPE (POTENTIAL).
 FT POLY-PRO.
 FT PDZ.
 FT C2 DOMAIN 1.
 FT C2 DOMAIN 2.
 FT S -> SGNIGIRIVGKEIPGSHGIGAYIAKILPGSGAE
 FT QTGKLMG (in isoform 2).
 FT /Frid=VSP 003923.
 FT K -> KPDTGKVVSHITGEIQ (in isoform 2).
 FT /Frid=VSP 003924.
 FT G -> GGVVTVQNAS (in isoform 2).
 FT /Frid=VSP 003925.
 FT TAHSK -> SKRRK (in isoform 2).
 FT /Frid=VSP 003926.
 FT Missing (in isoform 2).
 FT /Frid=VSP 003927.
 FT SEQUENCE 5147 AA; 563537 MW; CDS084390498CD3C CR664;
 Query Match 4.1%; Score 448; DB 1; Length 5147;
 Best Local Similarity 19.4%; Pred. No. 9,2e-08;
 Matches 501; Conservative 303; Mismatches 887; Indels 890; Gaps 112;
 QY 4 GAERKOSTPG-----SLFLSPAPAPKKGSSSDS--VGEKLGAAANAAYGRT 51
 DB 871 GAPKQAPAPSPQPTSGQPKSTGQAPAPAPAKSIPIVKKETAPAAEKLEPPAEQAPTVKR 930
 QY 52 EBYRRRTMTDKDGRGAATTTTTHRRFRRSVICSNATALBGLPLS-LPQPSIPAA 110
 DB 931 TETEKPPPI-KDSKSLTAEP-----QKAVL-----PTKLEKSPKRESTCPICKTEIN 977
 QY 111 VPQAPPE-----PH-REETVATAT--QVAGQPPAAABGE 145
 DB 978 IGSXDPHFNFCTECKQVCMCGFNPTPLTENQGTQRAISQUGDIRKMPAPSGP-- 1035
 QY 146 QAVAGPASTVPSSTSKRPVSGPGLVSGKEPPAPAGSGGSAKEQERSQQDDIE 205
 DB 1036 --KSPMPVPTSSSQKTAVPQVQLV-----KKQDEVTEAEKYL 1076

QY 206 ELIETAVGK-----SNDGRFLKFDIEIGSGSFXTVYKGLDTETTVAVMGELODRKLT 258
 DB 1077 EYKKEKTLISWEKIPWVATTDQKQESKLEBKXAS-----ALQEKKPL 1117
 QY 259 KSERGRFKEEMLMKGLQHPNIVRFYDSWESVYKCKICVLTEIMTSGTLKTYLRFKY 318
 DB 1118 PEEKKLIPEEKIRSEEEKP-----LLEKKLPPEDK--LPEAKTASAP-----EE 1162
 QY 319 MKIKYLRWCROILKGLQPLATRTPPPIIRHDKCNIPITG-PTGS---VKIGDIGLAT 373
 DB 1163 QKHDLIKS--QVQALAEELERVAIPKTYQEKQQTCKMEGLPSGPQSLPKREDVTYTK 1219
 QY 374 IRRASPAKSVICTPPMAPPEMTEEKYKDSVDVYAGCMLEMATSEYPISECONMAOI-- 431
 DB 1220 IKEQP-----QPPCTAKPD--QEKEDDKSDTSSSQQPKSPQSLPTGTSDDGSSSLBE 1271
 QY 432 -----YRRTSGYKPAFPDKVAIPE-----VKETIEGCIKONKDERYSIKD 472
 DB 1272 IESLIPTDEKDLKGLKQDSFOESSPSSDLAKLESTVLSLEQAATLADK-SEKK 1330
 QY 473 LNHAFQFQETGVRYELAEDEDEKI-----AIKMLRIEDIKKLGKYNDAIEPSPD 527
 DB 1331 TQPH-----EVSPEQPKQEKQSLSETLITISEEIKESQERRKD--FKXD 1377
 QY 528 LERDVEDVAQEMKESGYCEGDHKTMAAIKDRVSLIKRERQQLVNEDEKK----- 582
 DB 1378 SQQDIPS-----SKDHKEKSEFVVD--ITTRBPYDVSSESSSESPVP 1420
 QY 583 -----KQESS-----LKQVEQS 596
 DB 1421 QRRRTSVSSSSSDEYKQEDSQSGGEEDFIRQIIEMSADEASGEDEFIRNQLKEI 1480
 QY 597 SASQTIKOLPSASTGIPTAS-----TTSASVSTOVE-----DEEPA-DHQQLQY 642
 DB 1481 SSTBEOKEHTGKKGKIKTAGKRRLTRKSSSTIDDAGRHSHWDEDAFPESPLKY 1540
 QY 643 QQPSIVLSDPTVDSQGGSVF-----TESRV 669
 DB 1541 RFTKSQSEELVYTGCGGLRFRFTIETINSTIADKYSASSQKTSLYFDEPELEMSLT 1600
 QY 670 SSQQTYSYGQHQAIVSTGVPHIPSTV-----QAOSOPGV 707
 DB 1601 DSEDSRGRGSSSLHSAFTPTGTSVSLDEDSDSHKKGSKQKARHRHGP 1660
 QY 708 YPSSVACQ-----SQQPSSS-----LT 728
 DB 1661 LPLTIIDSSSEELREHEELLKQEKORIEQQOKSSKSKXKDELARRRRPKT 1720
 QY 729 GVSSSQPTQHPQOQCGIQTAAPQQTQVYLSQSTSTSEATTAQPVQPAQVLPQVSA 788
 DB 1721 PPSNLSPIEDASPTHEELROAAEMEELRRSSCSEYSPSE--SDPEGFETISPEKILIVQK 1777
 QY 789 GKQSTQGVQVAPAEVAVANOPATOPTLASSVDASHVASGMSDGENVSSSGRHE 848
 DB 1778 VYKLPVAVSLYSPTEDESLQKESQ-----KAKSAEEMYE 1814
 QY 849 GRTRKRYKRSVRSRHEKTSRPKILNVSNKGRVVEQ-----LETHNRKRV-- 899
 DB 1815 EMHKKH-----KYKAPPAANDREVFKEPELYGMLIEDYIESLVE 1857
 QY 900 -TEKFIDLDG--NPEIATITVNDFTLAIERSFPDQVREILEKADENLSEVSEPE 955
 DB 1858 DTVNGSDVGLTRQEE-----ENGPMQKGRQK-----RLSEQIYDPP- 1898
 QY 956 GDQGLESLQGDVDFGFSGQKLEGEFPQPTPSSMPQOIGITSLITQVHSAGRFIVS 1015
 DB 1899 -----MOKITD-----LQKEFYELBSLHVQEDLVSS-----FL-- 1930
 QY 1016 PVPESTRLESKYVPSKITD--TVAASTAQSPGMLSHSASSLISQAFSEL--RRAQMT 1070
 DB 1931 -IPESH-----ELVDLGTWVSTBEERKL-----LDADAAYEELMKQOMQLT 1972
 QY 1071 EG--PNTAPNPFSHTGPT--FPVP-----PLISSTAGVTTAAATAAPVATSSPPDI 1120

```

Db 1973 PGSSTQAPBEDMESTMDPEMDASLTSSVLSGASLTSTSSATLSIPDKITQH-F 2031
Qy 1121 STSVIOSE-VTPTEEGIAGVATSTGVVT-----SGGLPIPEVSSPLSSVSI----- 1170
Db 2032 STBEIEDYVVDYREIQEIHAHSLIITYSEPPSSAATVPP-SPTPLSTSSVSVCTTD 2090
Qy 1171 -----TTPAVSISTSPSLQVPTSTSEIIVSSTAL-----YP-----SVVATSASAG 1215
Db 2091 SSSPITLITTYTEPVDMTKPEDESEISSSTYFPFGSIIDYREHLSVLDRAAPDG 2150
Qy 1216 GSTA-----TPGPPRAVVS-----QQAGSTTVGATLT 1244
Db 2151 RASADHVIYISDMASSIIESVPPREGPVADVSTDLLISEKDPVKKAKKETGAILLE 2210
Qy 1245 SVSTTSPSTASQSLSTSSSTPTLAETV-----VSAHSLKTHSSST-----GLAF 1236
Db 2211 VLE-----AYDKKELEAKRTKSSISEVTFHPSPSYALPMKEQLSTTYTSGETF 2262
Qy 1297 SLASPPSSSSPGA-GVSSYSIQPGGLHPLVPSVIASTPLPQAAGPTSTPLPQVPSIP 1355
Db 2263 GOEKPAQOLPSGSPSVSLPAKPPFPSSSLDISAQP-PPPPPPPPPPPPPPPPPP 2320
Qy 1356 PLVGPVAVNAVQOTLIHSOPRALHPQPTHCPEDVDSTQPKAPGIDIKITLKKLS 1415
Db 2321 PLPPP-----TSPKPTILPKKLT-----VASPVYATPLFDVAT----- 2356
Qy 1416 LFSHSSSGAQAASVLSSTLVISVTWPGIPTAVAPSKULTSTSTCLPNTMLPGTV 1475
Db 2357 -----LETTAVLRSN-GLPVTRIC-----TTPAP----- 2379
Qy 1476 ALPTPVATGQVSTPVSTTSGV-----KPGTASBPPLTKAPVLPVGTLPAGTLPEQ 1531
Db 2380 --PVP-----KPSISPSGLVFTHRP--EPKSPPIAPKPIY-QLPTTQKPTD 2423
Qy 1532 LPPEP-GPSLTQSGQPELDDAQRRLSPKXITVTA-----VGPVMAAPRAITAGTQ 1586
Db 2424 IHKPTGLSLTSLK-----TLN-----LVTSADYKLPSPSPSLPSHS-----NKSS 2464
Qy 1587 POKGVSQVKEGPVLAATSSGAGVFMGRFQVSVADAGQKEGKKNSDEAKSVHSESTSES 1646
Db 2465 PRFSKSLTEHYVITLPSERG-----PTSSAS 2493
Qy 1647 SVLSS--SSESTLYKPEPNTITIGISS-DVPESAKHTTASAKSDTQPTKV--GR 1699
Db 2494 QALTSWPLGSPKDLVSEPFVPEPVTAVEIPISGEQFYISGALQTSAPVTAAPS 2553
Qy 1700 FQVTTTANKVGRF--SVSKTEKIDTKKEGP-----VASPPMIEQAVLPAVIT 1748
Db 2554 FQAAPTS--VTQPLTTEVSKTE--VSATKSTASVGLSISITIP--EPLALDNI-- 2603
Qy 1749 KKEPELSEPHLNGSPSPEAFLSRDVGSGSPHSPHOLSKSLPSONLSQSLSNSF 1808
Db 2604 HKEKPYKEDGKQ-----LVGDVILRTVKEVKTIDKICIDLSASTMDYKQI 2653
Qy 1809 NSSYMSDNESIDEDDLKLELRRLDKHKEIQDLSQKHEIESLYTKLGKVPVAVII 1868
Db 2654 TANEVYGAOISAVOPSIINLSV-----TSLVTPVSLATEVTIF 2692
Qy 1869 PPAAPLGRRRPTKSKSSSSSLGKSPOLSGNSGQSAASVLAHQQTLPHPGNIP 1928
Db 2693 VVCT-----ASASVYTGES--LVGAHEAMTTPLOLTSTSKAEPEYRIP 2734
Qy 1929 ESGQNLQPLKPSBSDNLYSAFTSDGA-----ISVPSISAPQOGSTNTVAT-- 1979
Db 2735 -----SDQVFIAREEARFVLSIGTRAHAYTLATKVTYTPPGVNTNGMTDSTVSQCIDTG 2790
Qy 1980 -----VNSQAQAQAPRA-MTSSRKGTETDIDLKLV----- 2008
Db 2791 EYVDLSTKSHRTVVTMESTSSVMTKILEDEKPYDLTGRRAVCCDVYKLPFGRSCTA 2850
Qy 2009 -----DNWARDAMNLSGRG-----SKGHMNTYEGGMARK 2038

```

```

Db 2851 QOPATLPEDEFGYRDDHYQYDRSPGYRGIGCKKPSMSDTNLAEGHFFYKSN-AFD 2909
Qy 2039 FSAPQCLISMT-----SNLGSAP-----ISAASATSLGHFTKSMCP--Q 2078
Db 2910 YSEGTDFVAVDLTSGRVTGTEVWDYSKRTGYPYETRGVTSAGISTPQYSTARKMPPPG 2969
Qy 2079 QY---GFPATPFQMGWSTGEPAPQGLQGFQPVGTASLQNTNISMLQKISNPPGSIKT 2135
Db 2970 QYCVGSVLRSSNGVSVATPTSTF-----ATTQPSGISFT 3008
Qy 2136 T 2136
Db 3009 T 3009

RESULT 7
MAPB_HUMAN
ID MAPB_HUMAN STANDARD; PRT; 2468 AA.
AC P46821;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Microtubule-associated protein 1B (MAP 1B) [Contains: MAP1 light chain
DE LCI].
GN MAP1B.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Fetal brain.
RX MEDLINE=95104835; PubMed=7806212;
RA Lien H.L., Feener C., Fischbach N., Kunkel L.M.;
RT "Cloning of human microtubule-associated protein 1B and the
RT identification of a related gene on chromosome 15.";
RL Genomics 22:273-280(1994).
CC -!- FUNCTION: The function of brain MAPS is essentially unknown.
CC Phosphorylated MAP1B may play a role in the cytoskeletal changes
CC that accompany neurite extension. Possibly MAP1B binds to at least
CC two tubulin subunits in the polymer, and this bridging of subunits
CC might be involved in nucleating microtubule polymerization and in
CC stabilizing microtubules.
CC -!- SUBUNIT: 3 different light chains, LCI, LC2 and LC3, can associate
CC with MAP1A and MAP1B proteins.
CC -!- DOMAIN: Has a highly basic region with many copies of the sequence
CC KKKK and KKEI/V, repeated but not at fixed intervals, which is
CC responsible for the binding of MAP1B to microtubules.
CC -!- PTM: LCI is coexpressed with MAP1B. It is a polypeptide generated
CC from MAP1B by proteolytic processing. It is free to associate with
CC both MAP1A and MAP1B. It interacts with the amino-terminal region
CC of MAP1B (by similarity).
CC -!- SIMILARITY: TO MAP1A.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
CC EMBL; L06237; AAA18904.1; -
CC Genew; HGNC:6836; MAP1B.
CC MIM; 157129; -
CC GO; GO:0005875; C:Microtubule associated complex; TAS.
CC InterPro; IPR000102; MAP1B_nuexaxin.
CC Pfam; PF00414; MAP1B_nuexaxin; 10.
CC PROSITE; PS00230; MAP1B_NUAXIN; 6.
CC Microtubule; Repeat; Phosphorylation.
CC CHAIN ? 2468 MAP1 LIGHT CHAIN LCI.
CC REPEAT 1878 1894 MAP1B 1.
CC FT REPEAT 1895 1911 MAP1B 2.

```

FT REPEAT 1912 1928 MAP1B 3.
 FT REPEAT 1929 1945 MAP1B 4.
 FT REPEAT 1946 1962 MAP1B 5.
 FT REPEAT 1963 1979 MAP1B 6.
 FT REPEAT 1997 2013 MAP1B 7.
 FT REPEAT 2014 2030 MAP1B 8.
 FT REPEAT 2031 2047 MAP1B 9.
 FT REPEAT 2048 2064 MAP1B 10.
 FT DOMAIN 589 790
 LYS-RICH (HIGHLY BASIC, CONTAINS MANY
 KEE AND KKEI/V REPEATS).
 SC SEQUENCE 2468 AA; 270618 MW; 540839CDBF09D461 CRC64;
 Query Match 4.1%; Score 439; DB 1; Length 2468;
 Best Local Similarity 20.6%; Pred. No. 8.4e-08;
 Matches 452; Conservative 271; Mismatches 828; Indels 640; Gaps 102;
 QY 90 ATALEPLGLPLSLPOPSI-----PAA--VPOSAPPEPREETVATATS 131
 DB 516 ATQKDLTG---QVPTVVVQTKLKQRADRESILKPAKPLPSKSVKESKEETPEVTKN 572
 QY 132 QVAQCPAAAARCEQAVAPAP---STVSSSTKDPVSGPSLVSGKEPPPARSGG-- 186
 DB 573 HV-EKPPKVESKKVWKDKDPVKTETKSVTEKEVP-----SKEEPSVKAIAVEK 623
 QY 187 ---GGSAAKE---POERSQOQDDIEELE--TKAVGMSNDGRFLKFD----- 224
 DB 624 QATDVPRKAKETVKTETKVPEDKKEKEKPKKVAKKEDKTIKKEKPKKEVKKE 683
 QY 225 --LEIGSGFKTYKGLDTEETVEVAMCEL--QDRKLTSEKRFYEBAAMLKGLQHPNI 280
 DB 684 VKKEIKKEKKEPKKEVKETPPKEVKEVKEKKEKKEKKEKKEKKEKKEKKEKKE 735
 QY 281 VREYDSMESIVKGGKCLVLTMLTSGTLKTYL--KRFKMKIKLRSWROLKGLQFL 338
 DB 736 -----PKDAKSSSTPLSEAKKPAALPKVPKKEESVKOSVAAKPKKE--KG--- 780
 QY 339 HTRPPIIHDLCNFI---TGPTGSYKIGDLATL--KRAFAKSVIGTEPFMA 391
 DB 781 ---KIVYKKEGKRAEVAAVAGTATTAVMAAGIAAGPAKELAEBSLMSPPDLT 837
 QY 392 PEMYEETKDESVDYAFGMCMLEMAITSEYSECQ---NAAQIVRTVSGVPRASFDKY 447
 DB 838 KD-PEELKAEVDVTKQIKPQLLEIIDEKLIKETEPEAVYIQRERVTKG--PAESPDE 894
 QY 448 AI-----PEYKEIIEGCIKONKDERYSIKLNLHAPFOE--ETGVAVELAE-- 491
 DB 895 GILTTEGEGCEQTPPELEPEVE--KQGVDDIEKFD--EGAGFESESSEGGDYEEKATE 949
 QY 492 ---EDDGEKIAIKMLRIEDIKKLGKXYKDNENAIFFSFDLEEDVEDVAQEWESGYV 546
 DB 950 EAEPPEEDGE-----EHVCVSASKISPTED--EESAKAEADA---YIREKRES--V 993
 QY 547 CEGDHKT---MAKAIKDRVSLIKRKERORLVREEOK---KKOESSILKQVQ--- 595
 DB 994 ASGDRAEDMDIAIE-----KGEAEQSEBEADEEDKAEDAREEVEPEMEAEEDYMA 1047
 QY 596 --SSASQGTG-----IKOLPSASTG-----IPTASTTASVSTQVPE 630
 DB 1048 VVDRAAAGAGAEBOYGLTTPYTKQIGASPGERPASSIHDETLGSGSESEKTA---DE 1103
 QY 631 EPADQHQOQL---YOOPSTISVLSDTG-----VDSGGSSSVFTESRVSQQTVS 676
 DB 1104 ENREDQBEFTATSGYOSTIEISSEPTPMDEMSTPRDVMSDENNETES--PSQEVN 1161
 QY 677 YGSGHEQAHSTGTVFGHIPSTVQAOSQPHGYVPPSSVAQOQ---SQQSPSSSLTGVSS 732
 DB 1162 ITKTESSLYS-----OEYSKPADVTPLNGFSESGSKTATDCKDNASASTISP 1209
 QY 733 SQPIQHPOQOQIGQGTAPPOQT-----VOYLSQSTSEBATTACVPSQPPQVLPQ 785
 DB 1210 PSSMEEDKFSRSLARDAVCSEKASTLTIDKDISIAVSEKVSPEKSPSLSPSPSPLEK 1269
 QY 786 VSAGKOS-----TGQSVQVAPABPVAVAQOPATQPTT---LASVDSASHSDVAGSMDSGN 837

DB 1270 TPLGRSVNFSILTPNEIKVSAEAEVAPVSPVETQVVEEHCASPDKLTLEVSP-----S 1324
 QY 838 ENVPSSSGRHGCRTRKRYKRSVRSRSRHEKTSREKRLILNVNSKGRVVECOLETHNKK 897
 DB 1325 QSVTSAGH-----TPYQSPTDKSSHLPTRE-----VIEKPPAV 1359
 QY 898 WTEFKEDLDGNPEBIATIMVNNDFILAIERESFVDQVREIIEKADMEISDVSEVEED 957
 DB 1360 PVSFEFS--DADONENRASVSPWDEPV--PDSESPYEKVLSPLRSPPLIGSE----- 1407
 QY 958 QGLSELQKDDYGFEGSGQKLEGEFQPLPASMPOQIGTPSSLTQV---HSGRRFI 1013
 DB 1408 SAVESFLSADDK--ASGRGAESPFEKSGKQSGPPQVVS--PVSEMTSTSLYQDKQEGKSTD 1464
 QY 1014 VSPVESRNRRESKVPSEETDTVAASTASQSPGMNLSH-----SASSLSLQO--AFSELR 1065
 DB 1465 PAPIKEDFGQEEK---TDVEMASSQ--PALALDERKLGQVSPQLIDVSGFGSKEDT 1517
 QY 1066 RAQMTGEPNTAPPNESHGTPPEVPPFLSSIAGVPTTAATAPVAPATSSPPNDISTSVI 1125
 DB 1518 KMSISEG-----TV 1526
 QY 1126 QSEVTVTEEGIA-----GYAT--STGVVNSGULPIPVGE--SPVLSVSVSTIPAV 1175
 DB 1527 SDKSATPVDEGVAEDTYSHMEGVA SVTASVATSSFPETTTDDVPSLHAIVGS----- 1580
 QY 1176 VGISTSPLOYPTSTSEIYVSTALYPSVTVATSSASAGSTATPGPRPVVSGQAAG 1235
 DB 1581 ---PHSTVEDLSISVAVQPTTFQETEMSPKSEEC-----PRP----- 1616
 QY 1236 STTVGATLTVSTSTTSPSTASQSLQSSSTPTPLAETVVVAHSLDKTSHSSTTGIA 1295
 DB 1617 ---MSISPPDPSPKYA-----KSRTPV-----QDRSEQSSMSIE 1648
 QY 1296 FSLSAPSSSSSPAGVSSITISQPGGLHPLVIPSIVASTPILPQAAGPTSPILLPQ----- 1350
 DB 1649 FGQESPEQS-----LAMDFSRQSPDHPVTVGAGVLIHT-----ENGETEVDYPSDMQDS 1697
 QY 1351 --VPSIPLVOPVAVNPVAVQOTLIHSOPALLPNQPHTCPEVSDTOPKAPGIDIKT 1408
 DB 1698 SLSHKIPMEHBPYTOQNDLISELI--SVSQVEASPSSTSAHTP-----SQIASPLQED--T 1749
 QY 1409 IBEKRLSLFSEHSSGAQAHASVLETSVLVETVTPGIPTTAVAPSKLTLSTSTCLPPT 1468
 DB 1750 L-----SDVAPPRDMSLYASTSR-----KVOSLBEKSLSPKS 1782
 QY 1469 NPLUGVALPVTVVTPPGOVSTSVSTTSGVKRGT---PSKPELTKAAPVLVGTGL-- 1522
 DB 1783 DIS-----PLTPRESPLVPTFSSTSAVKERTATCHSSSSSPPIDAASAEYGFRAV 1836
 QY 1523 -----PAGTLPSEROLPPFPGPSLTOSQOPLIEDLDAQLRRTLS 1559
 DB 1837 LPTDMQHLLANDLSTPGLEKXSGKTPDPFSYAAYQKPELTRSPDELDYDE-----S 1891
 QY 1560 PEKITVTSAVGPVSMAPTAITAGTOPQKVSQVKEGYPVLATSSGAGVTRMGHFOVVA 1619
 DB 1892 YEKTRTSVDG-----GYUYEK--IERTTKSP---SDSGSYETELGK--TYKTP 1933
 QY 1620 ADG-----AQKEGKNSDEAKSVHFPSSSTSESIVSSSSPEST---LVKPEPNC-- 1665
 DB 1934 EDGYSVYEIIEKTRTPEEGYSYDISEKTSPEVSGYETEKTERRRLLIDISNGYDD 1993
 QY 1666 -----ITIPGISVDVESAHKTTASEANS--DTGQPTKVGRFOVTT-----TANKYGRF 1712
 DB 1994 SEDGHTLGPBSYVETTEKITSFPEBEGSVYETS--TKTRTPTDSTYCYETAKITRT 2051
 QY 1713 SVS-----KTEDKITDTKKGVPVAPPPMDLQAVLPAVLPKPKKPELSESHLNGSSSD 1767
 DB 2052 PQASTYVETSDLCYTLEKKSPP--SEARQVDLCLVSSCEYKHKTKLS--PSFIN---PN 2105
 QY 1768 PEAFLSRVDVDDGSGSP-----HSPHQSSSLSPSONLSQSLSN 1806

Db 2106 PLEWFASEEPTESEKPLTOSGAPPPGKQOGRQCDTEPTVSSESAPQSDVDPPE 2165
 QY 1807 SPNSSYMSDNEDEDEDELE-RRURDKL-----KEIDDLGRQGH-----ETI 1853
 Db 2166 TEBCRSITADANIDSEDESRITPDKTITVYKMDPPAPVQDRSPRPHVPSWVDPAL 2225
 QY 1854 SLVYKLGVPVAVIIPPAAPLISGRRRRPTKSGKSSSSLSGNKSPQLSGNLGSGSAAAS 1913
 Db 2226 AIEONLGR-----ALKKDLKEXKTKKPGTK-TSSSPVYKS-----DQSKPEL 2268
 QY 1914 VLHPQQLTHPPGNIPESG-----QNLQLOPLKPSSSDNLYSAFTSDGAIYVS 1962
 Db 2269 AASPK-----PAGLKSSDDKVSRYVAPPKKESVKAAPTTTPVKARAGEKDKETGNA 2323
 QY 1963 LSAPGOGTSTNTVG-ATVNSQAQAOPPPAM 1992
 Db 2324 ANASAKSASAKATATAGPPTTKTKSSAVDPGL 2354

RESULT 8
 NEKA MOUSE
 ID NEKA MOUSE STANDARD; PRT; 792 AA.
 AC Q921J2; O35673; Q9R1J1;
 DT 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DE 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Serine/threonine-protein kinase Nek4 (EC 2.7.1.37) (NIMA-related protein kinase 4) (Serine/threonine protein kinase 2).
 GN NEK4 OR STK2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
 RC STRAIN=BALB/C; TISSUE=Spleen;
 RX MEDLINE=20001940; Pubmed=10529384;
 RA Hayashi K., Igarashi H., Ogawa M., Sakaguchi N.;
 RT "Activity and substrate specificity of the murine STK2
 RT serine/threonine kinase that is structurally related to the mitotic
 RT regulator protein NIMA of *Aspergillus nidulans*.";
 RL Biochem. Biophys. Res. Commun. 264:449-456 (1999).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RX MEDLINE=99321807; Pubmed=10393247;
 RA Chen A., Yanai A., Arama E., Kilfin G., Motro B.;
 RT "NIMA-related kinases: isolation and characterization of murine nek3
 RT and nek4 cDNAs, and chromosomal localization of nek1, nek2 and nek3.";
 RL Gene 234:127-137 (1999).
 CC -1- FUNCTION: Seems to act exclusively upon threonine residues.
 CC -1- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
 CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
 CC -1- ALTERNATIVE PRODUCTS:
 CC Name=1; Synonyms=STK2L;
 CC isoId=Q921J2-1; Sequence=displayed;
 CC Name=2; Synonyms=STK2S;
 CC isoId=Q921J2-2; Sequence=VSP_007001;
 CC -1- TISSUE SPECIFICITY: Expressed ubiquitously among various organs
 CC and is upregulated in the testis.
 CC -1- SIMILARITY: Belongs to the Ser/Thr family of protein kinases. NIMA
 CC subfamily.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AJ223071; CAA11072.1; -
 CC EMBL; Y09234; CAA70436.1; -

DR EMBL; AF099067; AAD16287.1; -
 DR PIR; JC7122; JC7122.
 DR HSSP; Q63450; 1A06.
 DR MGD; MGI:1344404; Nek4.
 DR InterPro; IPR000719; Prot kinase.
 DR InterPro; IPR008271; Ser_Thr_kin_AS.
 DR InterPro; IPR002290; Ser_Thr_kin_AS.
 DR InterPro; IPR001245; Tyr_kinase.
 DR Pfam; PF00069; Pkinase; 1.
 DR PRINTS; PR00109; TYRKINASE.
 DR ProDom; PD000001; Prot kinase; 1.
 DR SMART; SMO0220; S_TKC_1.
 DR PROSITE; PS00107; PROTEIN KINASE_ATP; 1.
 DR PROSITE; PS00111; PROTEIN KINASE_DOM; 1.
 DR PROSITE; PS00108; PROTEIN KINASE_ST; 1.
 DR Transferrase; Serine/threonine-protein kinase; ATP-binding; Mitosis;
 KW Nuclear protein; Phosphorylation; Alternative splicing.
 FT DOMAIN 6 261
 FT NP_BIND 12 20
 FT BINDING 35 35
 FT ACT_SITE 131 131
 FT MOD_RES 165 165
 FT VARSP_LIC 456 503
 FT FT
 FT CONFLICT 499 499 G->R (IN REF. 2).
 FT SQ SEQUENCE 792 AA; 88994 MW; CF9187311C807AID CRC64;
 Query Match 4.0%; Score 431.5; DB 1; Length 792;
 Best local Similarity 21.3%; Pred. No. 4.5e-08;
 Matches 190; Conservative 126; Mismatches 310; Indels 265; Gaps 30;
 QY 227 IGGSKTYTKYKGLDTEYTVVAMCELODRKLTYSERQFEAEMLKGLCHPNIVRFYDS 286
 Db 12 VGRGSYGEVTLVYHRDDGQYVYIKNLINASSRERRAAEQEQLSQLGHPNIVTFYKES 71
 QY 287 WESTVYKGCIVYVTELMSTGTLTKYTKRF--VMKIVLRSRCRQLKGLQTLHRTTP 344
 Db 72 WEB---GDGLYIVWFGPCGGDYRLKQKQGLPESQVVEFVIMALQYTHKK--H 126
 QY 345 IHRDLKCNIFITGPTGSKIGDLGLATL--KRASFASVIGTPEFMAPEWEEK-YDE 401
 Db 127 ILHRDLKQNVFLT-RTNIIKVGDLGIARVLNHHGMASTLIGTPYMSBELFSNKPYN 185
 QY 402 SYDVYAFGCMATSEYRPSYSCQNAQIYRRVTSVGPAPSFDAIPEVKEIIEGCI 461
 Db 186 KSDVMAIGCVVEMATLKAFNAKMNNSLVYRIIEGLPP--MPKYSTELABLRTMS 243
 QY 462 ONKDERYSIKDILNNAFPQEEIGVAVELAEEDGEXIAIKMIRIDIKLKKKYKDNEA 521
 Db 244 RRPERRPSVRSILRQPIYKHH-----ISLPLEATKATSKNNVKN-- 283
 QY 522 IEFSPDLERDVEDVAQENVESGYCEGDHKTMAKIKDVRSLIKRRRQRLVREOEK 581
 Db 284 -----CDSRAPFYA-----AVYSKESNDVVIHYQ-P 310
 QY 582 KKQESSSLKQVQEGSSAQ-----TGIKQLPSA--STGIPLASTTSASVSTQVEDEBE 633
 Db 311 RSESSEALHWMEDEKCLQEKVPDGLPLRSPASLGHGHTKQDMNNTGSECAT----- 362
 QY 634 ADHQOQLYQOQPSISVL-----SDGTVDSEGQSSVYTERVSSQCVVSGSQHEQAH 685
 Db 363 -----ISRINDILPAERDSANAGVYVESQIQHVDADADEVDSQCSISQEKRLQGN 414
 QY 666 STGVPIGHPITVQAQSPHGVPPSSVAQGSQ-----QPSSSSLATGVS 731
 Db 415 T-----KSSDPGNLLPERRSSDGGGBESSELVKKPLPENKQKPKPDQDYTGII 462
 QY 732 SSQPIQHQQQGGIGQITPPQCTVQYSLSQSTSEATTAQPVSPQAQVLPQVPSAGKQ 791
 Db 463 ENQDSIHRSQPHSSEPSLSRQRRQKREQTASGTYSQ--FQELPPRLP----- 513
 QY 792 STQGVSAVAPAEVAVVAPQATQPTTLASVSDVSAHSDVASGKSDNENVPSSSGHNEGRT 851
 Db 792 STQGVSAVAPAEVAVVAPQATQPTTLASVSDVSAHSDVASGKSDNENVPSSSGHNEGRT 851

Db 514 SYPGIGKV-----DIIATQONDGNGQGPVACVNSST 546
 Qy 852 TRHRKRSVRSRSHREKTSRPLRLINNSKGDVVECOLETHNRKMTFKFDLDGDNPE 911
 Db 547 S-----STSAK-----DRP- 556
 Qy 912 EATITVNNDFILATERESFVDVEIEIKADEMLSEDEVSEPEBGDGLSLQKDDYGF 971
 Db 557 -----LSARR-----RLTKSGEEMLPSPGAVQ-RTPSAVEPLKPGEE----- 594
 Qy 972 SGGQKLEGEFKQPIPPASMPQOIGIFTSLLTVVSHAGR--RFLVSPVPSERL----- 1022
 Db 595 -----DQPIPAQRSSD-----CSITQNMHTLPREKEXRLMHGLSEDELSSSTSS 639
 Qy 1023 -----RESKVFSEITDYVAASTAGSPGMMLSHSASSLSLQCAFSELR 1065
 Db 640 TDKSDGDSRSGKSHTEKMDLVQMLT-QTLRLKAKESCDLQVLNPGSEFR 689
 RESULT 9
 MINT_MOUSE STANDARD; PRT; 3644 AA.
 ID MINT_MOUSE 062504; Q80TN9; Q99P54; Q9Q2W2;
 DT 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE MSX2-interacting protein (SMART/HDAC1 associated repressor protein).
 GN MINT OR SHARP OR KIAA0929.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NC NCB1_TaxID=10090;
 RN [1]
 RP SEQUENCE OF 49-3644 FROM N.A. (ISOFORM 1), FUNCTION, SUBCELLULAR
 RP LOCATION, TISSUE SPECIFICITY, DNA-BINDING, AND INTERACTION WITH MSX2.
 RC TISSUE=Testis;
 RX MEDLINE=93379811; PubMed=10451362;
 RA Newberry E.P., Latifi T., Towler D.A.;
 RT "The RRM domain of MINT, a novel msx2 binding protein, recognizes and
 RT regulates the rat osteocalcin promoter.";
 RL Biochemistry 38:10678-10690(1999).
 RN [2]
 RP SEQUENCE OF 1-112 FROM N.A.
 RA Adachi J., Aizawa K., Akimura T., Arakawa T., Carninci P., Fukuda S.,
 RA Hachikume W., Hayashida K., Hirozane T., Hori F., Imotani K., Konno H.,
 RA Ishii Y., Itoh M., Kagawa I., Kawai J., Kojima Y., Kondo S., Kono H.,
 RA Koya S., Miyazaki A., Murata M., Nakamura M., Nomura K., Numazaki R.,
 RA Ohno M., Ohsato N., Saito R., Sakazume N., Sano H., Sasaki D.,
 RA Sato K., Shibata K., Shiraki T., Tagami M., Takeda Y., Waki K.,
 RA Watanuki A., Muramatsu M., Hayashizaki Y.;
 RL Submitted (OCT-2002) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE OF 69-3644 FROM N.A. (ISOFORM 2), AND VARIANTS THR-348;
 RP PHE-762; PHE-773 AND LEU-933.
 RC STRAIN=ICR; TISSUE=Brain;
 RA Sakamoto T., Gotou T., Isagawa Y., Mimura H., Kimura K., Kawaichi M.;
 RT "MINT/spen negatively regulates Notch signaling by inhibiting RBP-
 RT J/5u(H) activity.";
 RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE OF 318-578 FROM N.A.
 RC TISSUE=Cochlea;
 RX MEDLINE=97237053; PubMed=9119401;
 RA Crozet F., El-Amraoui A., Blanchard S., Lenoir M., Ripoll C., Vago P.,
 RA Hanel C., Fiammes C., Levi-Acobas F., Depertis D., Mattei M.-G.,
 RA Weill D., Pujol R., Petit C.;
 RT "Cloning of the genes encoding two murine and human cochlear
 RT unconventional type I myosins.";
 RL Genomics 40:332-341(1997).
 RN [5]
 RP SEQUENCE OF 2598-3644 FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=22579291; PubMed=12693553;

RA Okazaki N., Kikuno R., Ohara R., Inamoto S., Aizawa H., Yuasa S.,
 RA Nakajima D., Nagase T., Ohara O., Koga H.;
 RT "Prediction of the coding sequences of mouse homologues of KIAA gene:
 RT 11. The complete nucleotide sequences of 400 mouse KIAA-homologous
 RT cDNAs identified by screening of terminal sequences of cDNA clones
 RT randomly sampled from size-fractionated libraries.";
 RL DNA Res. 10:35-46(2003).
 RN [6]
 RP TISSUE SPECIFICITY;
 RX MEDLINE=22261914; PubMed=12374742;
 RA Oswald F., Kostezka U., Astrahantseff K., Bourteele S., Dillinger K.,
 RA Seidner U., Ludwig L., Wilda M., Hamelster H., Knoechel W., Lipay S.,
 RA Schmid R.M.;
 RT "SHARP is a novel component of the Notch/RBP-Jxappa signalling
 RT pathway.";
 RL EMBO J. 21:5417-5426(2002).
 RN [7]
 RP FUNCTION, AND TISSUE SPECIFICITY.
 RX MEDLINE=22483652; PubMed=12594956;
 RA Kuroda K., Han H., Tani S., Tanigaki K., Tun T., Furukawa T.,
 RA Taniguchi Y., Kurooka H., Hamada Y., Toyokuni S., Honjo T.;
 RT "Regulation of marginal zone B cell development by MINT, a suppressor
 RT of Notch/RBP-J signalling pathway.";
 RL Immunity 18:301-312(2003).
 CC -1- FUNCTION: Essential corepressor protein, which probably regulates
 CC different key pathways such as the Notch pathway. Negative
 CC regulator of the Notch pathway via its interaction with RBP5UH,
 CC which prevents the association between NOTCH1 and RBP5UH, and
 CC therefore suppresses the transactivation activity of Notch
 CC signaling. Blocks the differentiation of precursor B cells into
 CC marginal zone B cells. Probably represses transcription via the
 CC recruitment of large complexes containing histone deacetylase
 CC proteins. May bind both to DNA and RNA.
 CC -1- SUBUNIT: Interacts with NCOR2, HDAC1, HDAC2, RBBP4, MBD3 and
 CC MTAIL1. Interacts with the nuclear receptors RAR and PPARG.
 CC Interacts with RAR in absence of ligand. Bind to the steroid
 CC receptor RNA coactivator SRA (by similarity). Interacts with MSX2.
 CC Interaction between RBP5UH and NOTCH1.
 CC -1- SUBCELLULAR LOCATION: Nuclear. Associates with chromatin.
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=1;
 CC IsoId=Q62504-1; Sequence=Displayed;
 CC Name=2;
 CC IsoId=Q62504-2; Sequence=VSP 008564;
 CC Note=No experimental confirmation available;
 CC -1- TISSUE SPECIFICITY: Highly expressed in testis. Expressed at lower
 CC level in brain, lung, spleen, liver and kidney. Weakly expressed
 CC in cardiac and skeletal muscles and ovary. In spleen, it is
 CC expressed in follicular B-cells, while it is weakly expressed in
 CC marginal zone B-cells.
 CC -1- DOMAIN: The RID domain mediates the interaction with nuclear
 CC receptors.
 CC -1- DOMAIN: The SPOC domain, which mediates the interaction with
 CC NCOR2, is essential for the repressive activity (By similarity).
 CC -1- SIMILARITY: Belongs to the Spen family.
 CC -1- SIMILARITY: Contains 1 RID (receptor interacting) domain.
 CC -1- SIMILARITY: Contains 4 RNA recognition motif (RRM) domains.
 CC -1- SIMILARITY: Contains 1 SPOC domain.
 CC -1- CAUTION: Ref.4 sequence differs from that shown due to multiple
 CC framehifts and conflicts that create stop codons.
 CC -1- CAUTION: Ref.5 sequence differs from that shown due to what seems
 CC to be the presence of intronic sequence in the cDNA.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).

DR EMBL: BY726481; -; NOT_ANNOTATED_CDS.
 DR EMBL: AF156529; AAD55931.1; ALT_INIT.
 DR EMBL: AB055980; BAB32786.1; -.
 DR EMBL: Z78160; CAB01562.1; ALT_SEQ.
 DR EMBL: AK122402; BAC65684.2; ALT_SEQ.
 DR MGD: MGI:1891706; MInt.
 DR InterPro: IPR00504; RNA_rec_mot.
 DR Pfam: PF00076; xtm; 3.
 DR SMART: SM00360; RRM; 3.
 DR PROSITE: PS50102; RRM; 4.
 DR PROSITE: PS00030; RRM_RNP_1; FALSE_NEG.
 DR PROSITE: PS50917; SPOC; 1.
 KW Transcription regulation; Repressor; Nuclear protein; DNA-binding;
 KW RNA-binding; Repeat; Coiled coil; Alternative splicing; Polymorphism.
 FT DOMAIN 1 574
 FT 6 81
 FT 336 416 RNA-BINDING (RRM) 1.
 FT 439 514 RNA-BINDING (RRM) 2.
 FT 518 590 RNA-BINDING (RRM) 3.
 FT 559 575 RNA-BINDING (RRM) 4.
 FT 822 850 COILED COIL (POTENTIAL).
 FT 1185 1206 COILED COIL (POTENTIAL).
 FT 1509 1544 COILED COIL (POTENTIAL).
 FT 1607 1627 COILED COIL (POTENTIAL).
 FT 2216 2704 COILED COIL (POTENTIAL).
 FT 3478 3644 RID.
 FT 2138 2462 SPOC.
 FT 2706 2845 INTERACTION WITH MSX2.
 FT 125 277 INTERACTION WITH RBPSUH.
 FT 236 326 ARG-RICH.
 FT 648 721 SER-RICH.
 FT 702 832 TYR-RICH.
 FT 2101 2233 ARG-RICH.
 FT 2377 2518 ALA-RICH.
 FT 2950 3475 PRO-RICH.
 FT 618 640 Missing (in isoform 2).
 FT 348 348 /ftid=vsp_008564.
 FT 762 762 I -> T.
 FT 773 773 S -> F.
 FT 933 933 S -> F.
 FT 754 754 S -> L.
 FT 1524 1524 R -> G (IN REF. 3).
 FT 1560 1560 H -> A (IN REF. 3).
 FT 1570 1570 D -> Y (IN REF. 3).
 FT 1574 1574 F -> L (IN REF. 3).
 FT 1609 1609 R -> G (IN REF. 3).
 FT 1659 1659 Q -> R (IN REF. 3).
 FT 1669 1669 I -> V (IN REF. 3).
 FT 1705 1705 S -> F (IN REF. 3).
 FT 1815 1815 V -> A (IN REF. 3).
 FT 2097 2097 A -> V (IN REF. 3).
 FT 2201 2202 G -> A (IN REF. 3).
 FT 2322 2322 MISSING (IN REF. 3).
 FT 2385 2385 A -> V (IN REF. 3).
 FT 2502 2502 P -> Q (IN REF. 3).
 FT 2505 2505 R -> Q (IN REF. 3).
 FT 2519 2519 E -> K (IN REF. 3).
 FT 2554 2554 D -> N (IN REF. 3).
 FT 2679 2688 T -> S (IN REF. 3).
 FT 3010 3010 LSTPGAPVN -> VGEHWMARD (IN REF. 3).
 FT 3644 3644 L -> P (IN REF. 3 AND 5).
 SQ SEQUENCE 3644 AA; 398750 MM; 9C7EC49A81AD4A CRC64;
 Query Match 4.0%; Score 429.5; DB 1; Length 3644;
 Best Local Similarity 19.4%; Pred. No. 2.7e-07;
 Matches 443; Conservative 286; Mismatches 910; Indels 649; Gaps 98;
 Db 62 DKDSGAATTTTTHRRFRRSVIT-CDSNATALEPGLPLSPPSIPAAVQSAAPPREH 120
 1037 DTSRDGVLIRBESERKVRKEILRKESKTYLRLNSALSPKDCQDPAAVSAAGSGRPS 1096
 Qy 121 RE-----ETVATITSGVAQOPPAAPAGDAVAGPAPSTVP--SSTSKRPVSQPSLV 172

Db 1097 SDVHAGLCELTGHSVETQETO--PKKALPSK-----POPKOLQLLENQPEKEEVRKNYC 1149
 Qy 173 GSKERPPPARSGS---GGGSAKEPPEERSQOQDDIEETELKAVGMSNDGRFLKPIEIGR 229
 Db 1150 RPREEPARHRAQGEKPHGNAEE-----KGLDIDHQSTRKQMEGRKQRMETAK 1203
 Qy 230 G-SPTVYTKGLD-----TETT-----VEVAMELQDRKLTGS 260
 Db 1204 AEKFGSPKKDDVDYERSLVHVGKPPQDVYDDSPSKRRRTDHDVDFICTKREKNY-RS 1262
 Qy 261 ERQRFKEAEML-----KELQHPNIV----- 281
 Db 1263 SRQ-ISEDSERTSGSPVRHSGFHDDDDPGSPPLVSVKSGPKGDEKGLPYNAVARDP 1321
 Qy 282 ---RFDYSWESTVYKGCIVLVTLMTSGTLTKYLFKFKIKILRS-----WCQIL 332
 Db 1322 LKCNPDYS-----GKR-----EGTADATK-----IKLVNSEGSPRW----- 1355
 Qy 333 KGLQFLHTRTPPIIHRDLKCDNITITGPTGSKIGDGLATLTKRASPASKSVIGTPEPMAP 392
 Db 1356 -----DPEMKQDSRD--VSFPN-----SVTKDSLRKRSVDLE--P 1390
 Qy 393 EWEKEDSVVDVYAFGCMLEMATSEYPSQCNAAQIV----- 432
 Db 1391 GEVPSDDDEDAE-----HRSQPRASFTDSPPLSLFLDRDQKLERD 1434
 Qy 433 RRYTSGYKPAASF-----DKVAIPVKEIIEGC--IRONKDERYSIKDL-LNHAFQCEETG 484
 Db 1435 ERLASSLERNKFFSFALDKTITPDTKALERRAKSLSSRENNFSFLWDCRFANFRNNKD 1494
 Qy 485 VRVELAEEDDEGKIKALKLMRIEDIKLK-----GKYKNEALERSFDLERVDPEVAQ 538
 Db 1495 -----KAKDASAPRLPSWY---MKKKLRTDSBGLDDPKD-----ERREEOERQ 1538
 Qy 539 EWEESGYVEGDKHTAKAIKDRVSLIKRKEQ----- 571
 Db 1539 ELFASRFL---HSLTFEQDSKRLQHLERKSESDPPRGLYGRQMSGANSTSDSVQEP 1594
 Qy 572 -----RQVREDEQEK-----KQESSIKQVEQGSASQGTGKOLPSASTGI 613
 Db 1595 VLFHSRFMEITMOQKEKEKQPKAEKQKEPEETHPKPEPAALTKKEPEKAPASAGL 1654
 Qy 614 PTAST-----SASVSTVE-----PEEP-EADQHQOQOYQOQSIYSTLDGT 654
 Db 1655 PAVTIVTTPPEPSSAPEKAEAEAPASPAGEKAPAPVSEETKLVSEVPVAPQPRQ 1714
 Qy 655 VDSGQSSVFTESSVSSQOTVSYSQHOAHSGTGWPHIPISTVYQAOQSHGYPPSSVA 714
 Db 1715 SDVPPE-----DSRDSQDSALALAPAPQESAATDAVP-----CVNAEP---LTPGTYS 1761
 Qy 715 QGSGQGPSSSLTGVSSQPT-----QHPQOQGTQAPPQOQTVQVSLQSTSSBAT 769
 Db 1762 QVESSVDPEP-----SSPQLSKLTQRESEBAEG--KVEKDTTSTPEPATQNGVAS 1813
 Qy 770 TAQPVQAPQVLPQVSA-----GKSTQGVSAVAAPVAVAQAPQATOPTTIA 819
 Db 1814 EAQPAASEV--EAMPVAAKDKTKKSKRSKTSVQAAAASVVEKPTKRSERIDREKLR 1872
 Qy 820 SSVDAHSDVASGMDGEMNVSSSGRHEGRTTK-----RHYRKVRS---RSRH 866
 Db 1873 SSSPRGEAQKLELMMEBEKIRTASKSGSGDTEHPBPLPSRSRRNRVRSVYATMTIH 1932
 Qy 867 EKTSPKRLILNVSKGDRVVECCQLETHNRKNVTEKFPDLIDGNPEIATIMVNDPILAI 926
 Db 1933 ESRSPAKBPV-----EOPRVTEKRLERLEQEAIV-----TP----- 1965
 Qy 927 ERESFVDQVRELIEKADMLSEDSVVEPEQGLESLQKDDYVSGSGQKIEGFKQPIP 986
 Db 1966 PRRGPRPTTKRAEEDGHEHKEPEETPRPAAGWSPSSQSAAGQGGKRGNEQKVE 2025
 Qy 987 ASSMPQQIGIPITSLITQVHSAGRFIVSPVESLRSS-----KVPSSEI----- 1032
 Db 2026 AAA---EAGAGASTREGNPKSGREMASEPRKDRDRDPTDTSKSGPDTPVVEVLKRPPEK 2082

QY 1033 TDTVAASSTASPGMNLSSHASLSLQOAFSELRRAQMTSE-GENTAPPNFSHTPTFPVYP 1091
 Db 2083 TTKSKRGKRRKSTRSGMDRAAHQBSLEMAAQAQADKRAAASPOSESPOKSGSS 2142
 QY 1092 PRLSSAGVPTTAATAAPVATSPNDISTSE-----VIOSEVTPTEBGA---GVA 1141
 Db 2143 POLANNPADPDBAEESASASTAPEGTQLARQIELEQAVONIALPEPSAAASKGVA 2202
 QY 1142 TSTGVVTS-----GLPIPPSESPVLSVSSIT-----IPAVSISTSP 1183
 Db 2203 TATATASEEPAPEHGHKPAHQASETELAAGISIIISDASGEPEFNSAGAPSPVPSQTPH 2262
 QY 1184 S-----LQVPTSTSEIVSSSTALVP---STVSATSAQ-----GSTA 1219
 Db 2263 REGMEFGLHEASGILETGTATSSAPQVSALDPPEGSADTKTRNGSDSVQEAQSGVA 2322
 QY 1220 TPGKPPAVVSOQAAGSTTVGATLTSVSTTSPPTASQSLQLSSTSTPTLAETVVS 1279
 Db 2323 EWTPEPKDKGRQKTRRRKRNANKVVALTETRASEMQ-----TQSESPAEEATPAAT 2376
 QY 1280 AHS-----LDKTSHTSTTGAFSLASBSSS-----SPQAGVSYISQP- 1318
 Db 2377 PEAPQEEKSEKPPSPAECTFPDSKTPPAESLSQENSAAEKTCKAPVLPALPLPSQPA 2436
 QY 1319 -----GGHPLVIVSVIASPTILPOAAGPTSTPLLP---QVPSIPPLVQPVANVPAVOQTILI 1372
 Db 2437 LMDDGQARKKVSITLSDPVTSPDSGIPPTPLVITAKLPPVLP-GGV----- 2488
 QY 1373 HSOPQAL-----PNOPTHCEVSDTQPKAGIDITLKEKLSLSEH---S 1421
 Db 2489 HQBPPEVWTMITROEPPRAQSTPSPALPDTK-----ASDMQTSSTLTKIMDPKXVS 2543
 QY 1422 SSGAQAASVLETSVLESTVTTGCIPTTANAP-----SKLITSTSTCLP-----PMTLPL 1472
 Db 2544 ATGVTSTSV---TTALAEVPSAPCLQEPAPPCDPKHPLEGVSAAVANADQASBPV 2600
 QY 1473 GTVALPTPVTPGQVST-----PVSTTSGVKQGTASKPPPLTAPVLPVTEI- 1522
 Db 2601 AADKEKAPVIAPIAKTISVIRMPVSIULENSQKTTAKAPQITLGVSLTGLVNVSLV 2660
 QY 1523 PAGTLPS-----EQLPPFPSPILTSQOQPLEDLDAOLRTLSPEXI---TVTSAVG 1570
 Db 2661 PVNALKGPVKGVA TLKGLVSTPAGPVNLKGFNVLTGPVNVLTTPVSAITGVNAPAG 2720
 QY 1571 PVSMA-----AFTATEAGTOR-QKG-----VSQVKEGPVLTATSGA 1606
 Db 2721 PVTAAACGVTATGTAATGAVTAPAKGKORASSNENSRPHPGMSVVIDRPA-DTSGGA 2779
 QY 1607 GV-FKMGKFOYSVAADGAKQKQK-----SEDAKSVHFESESTSESVLSSSSPES 1656
 Db 2780 GLRVNTSEGVVLLSYSGQKTGPPRIAKISQITPPAAMIDEFOQSISKSVKXADS---I 2836
 QY 1657 TLVPEPENGITIPGISSDVPESAHTTAESAKSQDTGPTKGRFQVTTANKVGRFSVK 1716
 Db 2837 TPTQSAKGPOTPEAFANV--ANSTVLTLQTYNASP-----VLSVKTDRPSLEK 2886
 QY 1717 TE-----DKITDTKEGPVA-----SPFMDLEQAVL--PAVITKEKEPELSESHANGP 1764
 Db 2887 PEPIHLASVITPVOGGTVKVLTOGINTPVPVWNLVLTLSIVTNNK--LADPVTLKIE 2944
 QY 1765 SSDEEAAFLSDVDGSGSPHSPHQLSKSLPSQNLQSLSNFSNYSMSDNESDIDDE 1824
 Db 2945 TKVLQPAVLGPTL-----TPHHPALPEK-LPA----- 2971
 QY 1825 DLKLELRRLDKHLEIQLOSROGHEIESLYTKLGKVPVAVILP-PAAPLGGRRRPTK 1883
 Db 2972 -----EVNHVPSGPETPADRTIAHLATPKDPTHSRPRGPTPGJFPRCH 3016
 QY 1884 SKGSKSSSSSLGKNSPOLSGNLSCQASVAVHPQOT-LHPPGNIPES-----GQNO 1934
 Db 3017 PSSSTTSTALST--NATVWLMAGIPIVPOFPISSIHQSVIMPHSITQTVSLGHSQGBVR 3074

QY 1935 LIQPLKPS 1942
 Db 3075 MSTPLPS 3082

RESULT 10
 ID PCLO RAT
 AC 09UKS6; Q9ULT1; STANDARD; PRT; 5085 AA.
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Piccolo protein (Multidomain presynaptic cytomatrix protein).
 GN Piccolo.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OC NCBI_TaxId=10116;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM 2), AND INTERACTION WITH RABAC1.
 RX MEDLINE=20170257; PubMed=10707984;
 RA Fenster S.D., Chung W.J., Zhai R., Cases-Langhoff C., Voss B.,
 RA Garner A.M., Kaempf U., Kindler S., Gundelfinger E.D., Garner C.C.;
 RT "Piccolo, a presynaptic zinc finger protein structurally related to
 RT bassoon.";
 RL Neuron 25:203-214(2000).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RA Fenster S.D., Cases-Langhoff C., Gundelfinger E.D., Garner C.C.;
 RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP CALICUM-BINDING ACTIVITY, AND MUTAGENESIS OF ASP-4668, ASP-4674;
 RP VAL-4688; MET-4689; VAL-4690; SER-4691; GLN-4692; ASN-4693 AND
 RP ALA-4694.
 RX MEDLINE=21181819; PubMed=11285225;
 RA Gerber S.H., Garcia J., Rizo J., Suedhof T.C.;
 RT "An unusual C(2)-domain in the active-zone protein piccolo:
 RT implications for Ca(2+)-regulation of neurotransmitter release.";
 RL EMBO J. 20:1605-1619(2001).
 CC -1- FUNCTION: May act as a scaffolding protein involved in the
 CC organization of synaptic active zones and in synaptic vesicle
 CC trafficking (By similarity).
 CC -1- SUBUNIT: Interacts with Rabac1/Pral, RIM2 and profilin (By
 CC similarity).
 CC -1- SUBCELLULAR LOCATION: Concentrated at presynaptic side of synaptic
 CC junctions.
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=1;
 CC IsoId=O9UKS6-1; Sequence=Displayed;
 CC Name=2;
 CC IsoId=O9UKS6-2; Sequence=VSP_003930, VSP_003931;
 CC -1- DOMAIN: C2 domain 1 is involved in binding calcium and
 CC phospholipids. Calcium binds with low affinity but with high
 CC specificity and induces a large conformational change.
 CC -1- SIMILARITY: Contains 2 C2 domains.
 CC -1- SIMILARITY: Contains 1 PDZ/DHR domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb.ch/announcements>;
 CC or send an email to license@isb.ch).
 CC -----
 CC EMBL; AF138789; AAF07822.2; -
 CC EMBL; AF227534; AAF63196.1; -
 CC HSSP; P04410; 1A25.
 CC GO; GO:0045202; C:synaptic junction; IDA.
 CC GO; GO:0005509; F:calcium ion binding; IDA.
 CC GO; GO:0005544; F:calcium-dependent phospholipid binding; IDA.
 CC GO; GO:0005522; F:profilin binding; ISS.

DR GO; GO:0007010; P:cytoskeleton organization and biogenesis; ISS.
 DR GO; GO:0016080; P:synaptic vesicle targeting; NAS.
 DR InterPro; IPR000008; C2.
 DR InterPro; IPR001478; PDZ.
 DR InterPro; IPR008899; Znf_piccolo.
 DR Pfam; PF00168; C2; 2.
 DR Pfam; PF00595; PDZ; 1.
 DR Pfam; PF05715; Zf_piccolo; 2.
 DR SMART; SM00239; C2; 2.
 DR SMART; SM00239; PDZ; 1.
 DR PROSITE; PS00499; C2_DOMAIN_1; 1.
 DR PROSITE; PS50004; C2_DOMAIN_2; 2.
 DR PROSITE; PS50106; PDZ; 1.
 DR Calcium/phospholipid-binding; Metal-binding; Zinc; zinc-finger;
 Repeat; Alternative splicing.
 KW DOMAIN 372 491
 FT 12 X 10 AA TANDEM APPROXIMATE REPEATS OF
 P-A-K-P-Q-P-Q-Q-P-X.
 FT C4-TYPE (POTENTIAL).
 FT C4-TYPE (POTENTIAL).
 FT POLY-PRO.
 FT PDZ.
 FT C2 DOMAIN 1.
 FT C2 DOMAIN 2.
 FT TKPTN -> SKRRK (in isoform 2).
 FT /FTID=VSP_003930.
 FT /FTID=VSP_003930.
 FT Missing (in isoform 2).
 FT D->A: COMPLETE LOSS OF CALCIUM-BINDING
 AND CALCIUM-DEPENDENT PHOSPHOLIPID
 BINDING ACTIVITY.
 FT D->A: COMPLETE LOSS OF CALCIUM-BINDING
 AND CALCIUM-DEPENDENT PHOSPHOLIPID
 BINDING ACTIVITY.
 FT V->S: SMALL INCREASE IN AFFINITY FOR
 CALCIUM.
 FT VM->SS: 10-FOLD INCREASE IN AFFINITY FOR
 CALCIUM.
 FT M->S: INCREASED AFFINITY FOR CALCIUM.
 FT VV->SS: 10-FOLD INCREASE IN AFFINITY FOR
 CALCIUM.
 FT ON->AA: MODERATE INCREASE IN AFFINITY FOR
 CALCIUM.
 FT CALCIUM.
 FT A->S: NO EFFECT ON CALCIUM-BINDING
 ACTIVITY.
 FT MUTAGEN 4674 4674
 FT 4674
 FT MUTAGEN 4688 4688
 FT 4688
 FT MUTAGEN 4689 4689
 FT 4689
 FT MUTAGEN 4690 4691
 FT 4691
 FT MUTAGEN 4692 4693
 FT 4693
 FT MUTAGEN 4694 4694
 FT 4694
 FT SEQUENCE 5085 AA; 552702 MM; 5A1BB543201A7450 CRC64;
 Query Match 3.9%; Score 423; DB 1; Length 5085;
 Best Local Similarity 19.9%; Pred. No. 6.3e-07;
 Matches 404; Conservative 270; Mismatches 757; Indels 596; Gaps 87;
 21 PAPAPNKGSSSDSSVGEKLGAAADAVTGRTEYRRRRHTMDKDSGAAT--TTTEHR 78
 1087 PMPAPAPESPQQKPTGTQVKGKKKA-EKTE--AEKVPKEKETSIEKTPMTTDDK 1142
 79 FRRRSVTCNSATALELPGIPLSLP-----OPSI-----PAVPOS 114
 1143 -----LESEEGKKSVALPEKKPEEEKKASADKERRKPAEEKRPLEKKPIPVCK 1196
 115 APPEPH-----REIVTATATSOVAQPP--AAAPGE-----QAVAPAPSTV 156
 1197 LPPEAKPLSSEGEKEKILKAHVQIPEEPPTGKVAKAGEEEOQDPSREALPGATPLTL 1256
 1257 PRAGEKERAVAPQAGSSKDDGERS-----KEKTEKEDKSDTSSEOQPKSGIS 1309
 157 PSTSKDRPVSQPSLVSGKEPPAPASGSGSAGKPEQERSQOQDDIELETKA----- 211
 157 PRAGEKERAVAPQAGSSKDDGERS-----KEKTEKEDKSDTSSEOQPKSGIS 1309
 212 -VGMNDGFLKFDLIGSGFKTYVKGIDTETVEVANCLEQDRKLTSEORFKEAE 270
 1310 DTGYSDDG-----ISGLGEIPLSLP-----SDEED 1335
 271 MKGKQHPPIVAFYVSWEST-----VKGKCIIVLVELMTSGTL-KTYLKRFPVMKIK 322
 1336 LKGLKKDSF-----SQESSPSSPDLALETSTVSLIEAOSTLVGEKAEKTKPOKIS 1390

QY 323 VLRSKQIILKGLFLATRPPIIHRDKCDNPIFGTSGVKIGDLIATIKRASFAS 382
 Db 1391 PEKFOQDQ-----KTQT-----ASPLDITISEEIKSGEKKSPK 1427
 QY 383 VIGPEPAPMEYKEDSDVYAFGCMLEMTSETPY-----SECNAQIYRYVT 436
 Db 1428 KDSQGFPSRKEHEK-PELVD-----DLSPRASYSVEDSESESPVRRRRR 1477
 QY 437 SGVPASFDKVALPEVEKIEIGCIRONKDERYSIKDLINAHFOEET-----VRVLA 490
 Db 1478 TSIGSSSD-----EYQ--EDSGSGEEDPIKQIITENSADEBASGEDEEFTSOLK 1530
 QY 491 E-EDDGKAIKMLRIDIKLKKYK-----DN 519
 Db 1531 EISGVGES-----QKREBAKGGKGVAGKRRILTRKSTSFDDAGRRHSHMEDD 1581
 QY 520 EALFEFDLE-RDVPEDVAQEMVESGYVCEGDNHT--MAKAIKRVSLIKRKRQOLVR 576
 Db 1582 ETPDESELPKRETKSGESELVAVGGGLRRFTIETLNSTIADKYS---SESSQKTTIL 1638
 QY 577 EDEKKKQESSSLKQVQSSASQGTGKLPAS--TGIPASTTSASVSTQVEEPPEA 634
 Db 1639 YPDEPEPELEMESLIDSPEDSRGE-GSSSLHASFPTGTSTVSILDEDSSESHKKG 1697
 QY 635 DQHQQLQYQPPSISVSDGTVDGSGSSVFTESRVSSQQTVSQSGQHQHASTGTVPCHI 694
 Db 1698 ESQQRKARHRSHGPLFTTEDSEEEELKEEELKEQKQRELEQOQKSSSKSKKD 1757
 QY 695 PSTVQAG-SQPHGVPPSSVAQGSQGSQSSSLTGVSQPIQHPPQOQGIQTAPQ 753
 Db 1758 KDELRQRRERKPTPSNLS-----PIEDASPTTELQAAEMEE 1797
 QY 754 TVQYSLSQTSSEATTAQFVSQPAQVLPQVSAQKSTQGSQVAPAPVAPQAT 813
 Db 1798 IHRSSGSEVSPSTE---SDPEGEISPEKILYQKXKILPAVSLYSFTDEQSVQKEGV 1854
 QY 814 QPTLLASSVDASHDVAAGSGDGENAVPSSGHEGRTTYRRYKRSRREKTSRPK 873
 Db 1855 Q-----KALKSAEYEEEMQKPH-----K 1874
 QY 874 LRLIYNSKQDRVVECO-----LETHNRKXV--TFKFDLGD--NPEIATIMVN 920
 Db 1875 YKAFPAANRDEVEFEKPYLGGMLIEDYIESVEDTYNGSVGSLTRQEE-----QN 1928
 QY 921 DFLALIERSPVDQVREIIEKADDEMLEDVSVPEBQGLSEILOGKDPYGS----- 972
 Db 1929 GFMOQGRGQKVRLEQIYDDMQKIS-DLOKEFYELSHSVRPQEDIVSSYIIPESH 1987
 QY 973 -----GSKLE-GERKQPLPAS-----MPOQIGIPTSLLQVHSAGRFTVSPVES 1020
 Db 1988 EIVDLSGVMYSTSEEEKLIDADSAVEELMRQOVYTDGS-----SPQYTGIDMAES 2041
 QY 1021 RLRESKVFPESEITDPAVASTAOSPQNLSHSAISLQ-----QAFS--ELRQOMTE 1071
 Db 2042 TLDPBRVQDASLTSLS-----QASLTDSTYSATLSIPDYKITQCFABELEDYVD 2095
 QY 1072 GNTAPNPFESHTGTPFPVPPFLSIAGVPTTAAATAPVATSPENDIS--TSVIGSEV 1129
 Db 2096 YTREIIDIILHBSLILTYSEPEES-----ATSVPESDIPSLTSSISVC 2139
 QY 1130 TVPTEGIAVATSGVTSGLPIPVSESVLSSV--SITTPAVAVIS--TSP 1183
 Db 2140 TTDSSSPYTTLDLITVYTEPADVMTKFDSEIESTYFPGSIIIDYEDISVLDRTIMP 2199
 QY 1184 SLQVPTSTSEIYVSTALYPSVTVSATSSAGSGATATGPKRPVAVSQAGSTVG-- 1240
 Db 2200 BSR--TNEDRIYLSGAPSVSVSGIKPBPQADTISTDPISEKDLIKKKEKTEGCI 2257
 QY 1241 -----ATLTSVTTSPSTASQSLIOLASSSTPTLAETVVASHSIL 1283
 Db 2258 ILEVLDAYKDRBESEALTKISLPE--PGA-----QAPSVTAFOIKQ--HVSHPHVS 2309

```

QY 1284 -----DKTSHSSTTGIALFSLAPSSSSPGAGVSS--YISQPGIHLPIVPSVIA-- 1332
DB 2310 GMSIGGEKFTYRLPSG-SLPVSTHPSKRPFRSSSLDSADPPPPPPPPSPSSSP 2368
QY 1333 -TPILPQAGP-----TSTPLIPQVPSIPPLVQVPA-----NVPVQOT 1370
DB 2369 PFPPLPAPSPKRPPTPKKLAVALATVSTTIVTHTVDALVWEAAARNSNGLPATKMC 2428
QY 1371 LHS--OQPALP-NQPHITCPV-----DSTQPKAPSDIDI 1406
DB 2429 AAPPVPKPKPSQIPGTGLVTHRPEAIKPIAPKPAVPOI PVTQKPTCTCPKPTGLSLT 2488
QY 1407 KTLSEEL-----SLFSEHSSGAGHSAVSL-ETSLVSESTVPGIPTTVA-- 1452
DB 2489 STWSLNLVSADYNVSPSPSPSPKNSPRKSLMTYVITLPSRPGPTDSSAAQ 2548
QY 1453 -----PSKLLTS-----TSTCLPPTNLP-----LGTVALPYTPVT-- 1484
DB 2549 AITSMPLGSPKOLVLETFVSVPPMTSTETPSAQPTLYTSGALGTPS--VTPAVTAS 2606
QY 1485 -----PGQVSTPVSTTSGVKGTPAKSP--PLTKAPV-----LP 1517
DB 2607 LFOVPTSLTQPLPAESKSEVSAVSAVSAVSAVSAVSAVSAVSAVSAVSAVSAVSAV 2666
QY 1518 -VGTLPAGTLPSSEOLPPFPSPSLTQSQPLE-----DLDAQ-----LRTTIS--PEXI 1563
DB 2667 LIGDAIDLRTIPKSEV-----KATCEKMLLSANDVKQOTANVYRQIQIAVQPSIT 2720
QY 1564 TVTSAVG--PVSM-----AAPTAITEAGTQPKQVSYQVKEGVLTSSGAGVFKK 1611
DB 2721 NLSAASLGTPTVMDSKTVAVVTCTDTTIVTGTESQVGEHVAVTSPLQLTTS----- 2773
QY 1612 GRRQVSAVAAGAQKGEKNSKEDAKSVHFSESTSESSVLTSSSPES--TLVKDEP----- 1663
DB 2774 -----KHTELPYRKPSSQAFTIRDEAPINLSLGSAAQVATVTKPVTVP 2820
QY 1664 -----NGTIPGISDVPE-----SAKTTASEKSDTGOPTKV 1697
DB 2821 VGVTNGMTDSTLSQVADGEVLDLSTSKSRITVVTMDSTSNVTXI 2867

RESULT 11
NRKI_YEAST STANDARD; PRT: 1080 AA.
ID NRKI_YEAST
AC P38692;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Serine/threonine-protein kinase NRK1 (EC 2.7.1.37) (N-ribose kinase 1).
GN NRK1 OR KIC1 OR YHR102W.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DC-5;
RA Fukumai Y.;
RL Submitted (MAY-1994) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=SG288C / AB972;
RC MEDLINE=94378003; PubMed=8091229;
RA Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Dover J.,
RA Du Z., Favello A., Fulton L., Gattung S., Giesel C., Kisten J.,
RA Kucera T., Hillier L.W., Jier M., Johnston L., Langston Y.,
RA Lareille P., Louis E.J., Macri C., Mardis E., Menezes S., Mouser L.,
RA Nguan M., Rifkin L., Riles L., St Peter H., Trevisan S., Vaughan K.,
RA Vignati D., Wilcox L., Woldman P., Waterston R., Wilson R.,
RA Vaudin M.;
RT "Complete nucleotide sequence of Saccharomyces cerevisiae chromosome
RT VIII."
RL Science 265:2077-2082 (1994).

```

```

CC -! CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC -! SIMILARITY: Belongs to the Ser/Thr family of protein kinases.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```

```

CC EMBL: D29980; BA06250.1; -.
CC EMBL: U00059; AB06860.1; -.
CC PIR: S48944; S48944.
CC Germonline: 139419; -.
CC SGD: S0001144; KIC1.
CC GO: GO:0016301; P:kinase activity; IDA.
CC GO: GO:0007047; P:cell wall organization and biogenesis; IGI.
CC GO: GO:0009092; P:cellular morphogenesis; IGI.
CC InterPro: IPR000719; Prot kinase.
CC InterPro: IPR008271; Ser Thr pkin AS.
CC InterPro: IPR002290; Ser thr kinase.
CC InterPro: IPR001245; Tyr_kinase.
CC Pfam: PF00069; PKinase; 1.
CC PRINTS: PR00109; TYRKINASE.
CC ProDom: PD000001; Prot_kinase; 1.
CC SMART: SM00220; S_TKc; 1.
CC PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
CC PROSITE: PS00108; PROTEIN_KINASE_SF; 1.
CC PROSITE: PS00111; PROTEIN_KINASE_DOM; 1.
CC TRANSFERASE: Serine/threonine-protein kinase; ATP-binding.
CC DOMAIN 23 276
CC FT NP_BIND 29 37 ATP (BY SIMILARITY).
CC FT BINDING 52 52 ATP (BY SIMILARITY).
CC FT ACT_SITE 144 144 BY SIMILARITY.
CC SQ SEQUENCE 1080 AA; 117061 MW; 99899EAF315E0B94 CRC64;

```

```

Query Match 3.8%; Score 416; DB 1; Length 1080;
Best Local Similarity 22.1%; Pred. No. 2.1e-07;
Matches 257; Conservative 173; Mismatches 473; Indels 262; Gaps 52;

```

```

QY 227 IGRGSEFTYVKGALDTETVEVANCELODRKLTSEKQFPEEEMKGL-QHNINIRFYD 285
DB 29 IGRGSEFTYVKGALDTETVEVANCELODRKLTSEKQFPEEEMKGL-QHNINIRFYD 87
QY 286 SMSSTYKGGKCCIVLNLVETLMSGTLKTYLKKFKYKIKVLSMCRQILKGLQFLHTPTPI 345
DB 88 SYLKDTG---LMTIMEHCGSLRSILRGKIDE-KYIGVIRRELVALKCHKON--V 140
QY 346 IHRDLKCDNIFITGPTGSYKIGDLGLA-----TLKRASPAKSVITGPEFMADPMTEE-- 397
DB 141 IHRDLKCDNIFITGPTGSYKIGDLGLA-----TLKRASPAKSVITGPEFMADPMTEE-- 195
QY 398 KYDESVDVYAFGCMLEMAISEYPSYSECQAAQIYKRVTSYGVCPASDKAIVEKKEIE 457
DB 196 KYDESVDVYAFGCMLEMAISEYPSYSECQAAQIYKRVTSYGVCPASDKAIVEKKEIE 254
QY 458 GCIKRONDERYSIKDLNHAFFQETGTVRVELAEDEDEKXIAIKLWIRIDIKKAKGYK 517
DB 255 LCLDEDEKXIAIKLWIRIDIKKAKGYK-----LISRYLPLPD--KNNKKYK 306
QY 518 DNEAIEBSPDLERDVPEDVAQEVESGYVCE--GDHKYAKAIKDRVSLIKRREBQRL 574
DB 307 -----IGSIPENEPSPKPEAPKPSQNGCGD-----EAKKSI 338
QY 575 VREQEKKKQDESSLKQOVSSASQGI-----KQIPASASTGIPASTTSASVSTQVEP 629
DB 339 VREQEKKKQDESSLKQOVSSASQGI-----KQIPASASTGIPASTTSASVSTQVEP 676
QY 630 E-----PEADHQOQ-QYQOQPSISVLDGTVDGQGSV-----FTSRVSSQOQTS 676
DB 399 DEDSYDPPSHMTNRPVYQCTTIGKGVPGI--AQNSTLNAVPTNNYTNKSPYKQVAG 456

```

```

QY 677 YGSGHEQAHSTGT-VPGHIPSTVQAQSQ-----PHGVYPP 710
Db 457 -----TNTSGTHAGPMTSKRLESKAPKQLLEFENETITANDVNTAPKISISI 510
QY 711 SSVAGOGGOG---PS-SSSLTG-VSSSOPIOH-PQOQOG---IQOTAPQOQTVGSLSQ 761
Db 511 SSLNAGNSRRDFFPISINENVGNNINNNKMRPDLPLSSGNNVYQSQTALPFLQTKFNK 570
QY 762 TS-----TSSEATTAQVPOAPQOVLPVSAKQSTQGVASQVAPAPVAVQAQAPQPTL 818
Db 571 TSKGPTSGLTTPATPISIELEIP-----NSALPTPASADPVL-----PSTK 615
QY 819 ASSVDSANSDVASCMDGENVPSSSGRHEGTTKRYKRSKRSRHE-----KTSRKL 874
Db 616 ARS-----STVYAG-----TPSSSSSIQYKSPENVPRRLTVSNRRHECPSTITNKL 663
QY 875 RILVNSNGKRDVCEQLETHRKRVTPFDLDGONPEIATIMVNDITALEBSFVDQ 934
Db 664 GSAVASNGISSTPNSNNNN-----NTDSENSRSGSSGNTANSTQMGITPQVTK 716
QY 935 VREIIEKADEMLSDVSVPEPDQGLSLQKDDYGFSGSKLGEFPKQPIPASMPQOI 994
Db 717 -----LSTHKASSPS-----RPLFG---VGTSPNK-----PAQSPTONI 748
QY 995 GIPTSSTLVVHSGKRFIVSPVESRLRESK---VFPSEITDTVAASTQSPGMILSH 1050
Db 749 G-----HNSHTNLAPPTMKPMANSKXNDKILLQPLNSIPSSSTLNTISGN--- 795
QY 1051 SASLSLQAQAFSELRRQMTGEPNTAPPNPHPTFPVPPFLSLAGVPTTAATAAPV 1110
Db 796 SSNNLTSSNFSNKESSRVNG-----DEKRNPNLKLQWPL-----PT 834
QY 1111 PATSSPNDISTSVIOGEVTPTEGIAVATST-----GVWTSGLPIPPVSPPLSS 1165
Db 835 PVVANKLLDPTANSGNNNGP---GSAGITNNINNOFGENISSASNT-PVSTTPISEK 890
QY 1166 VWS-SITTPAVVSTSTSPSLQVPTSTSEIYVSTALPSTVYVNTASAGSTAPPGPK 1224
Db 891 HIFPGKIKXSQSSINRKN-----SASEHPNLNLIGSSVSGVNGGICNNNVGNNNSGN 945
QY 1225 PPAVSSQAAGSTTVGATLTSVSTTSFPSTASQLSTQLSSSTSTPLATVVA---H 1281
Db 946 NSVPLSAN-TGNTTIRKANTSTIANSS-----AAASTTAPISQOTIPSGTOFHN 993
QY 1282 SLDK--TSHSSTGLAFSLAPSSS 1304
Db 994 ILSSAATVANSVNSIGCMCPPOQS 1018

```

RESULT 12

```

NC06_MOUSE STANDARD; PRT; 2067 AA.
AC 09JUL19; 09JUL19;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Nuclear receptor coactivator 6 (Amplified in breast cancer-3 protein)
DE signal cointegrator-2 (ASC-2) (Peroxisome proliferator-activated
DE receptor-interacting protein) (PPAR-interacting protein) (Nuclear
DE RAP250) (NRC) (Thyroid hormone receptor binding protein).
GN NC06 OR AIB3 OR RAP250 OR PRIP OR TRBP.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2), AND INTERACTION WITH
RP PPARG; PPARG; RARA; RXRA; ESR1; ESR2 AND THR.
RC TISSUE=Liver;
RC MEDLINE=20250907; PubMed=10788465;
RA Zhu Y.-J., Kan L., Qi C., Kanwar Y.S., Yeldandi A.V., Rao M.S.,

```

```

RA Reddy J.K.;
RT "Isolation and characterization of peroxisome proliferator-activated
RT receptor (PPAR) interacting protein (PRIP) as a coactivator for
RT PPARG."
RL J. Biol. Chem. 275:13510-13516(2000).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM 2).
RC TISSUE=Breast;
RC MEDLINE=22386257; PubMed=12477932;
RA Klausner R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shennen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Ditchenko L., Maruina K., Farmer A.A., Rubin G.M., Hong L.,
RA Brownstein M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.B.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McGowan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Huliy S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.W., Marra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [3]
RP SEQUENCE OF 786-1142 FROM N.A. (ISOFORM 1), INTERACTION WITH PPARG;
RP PPARG; ESR1; ESR2; THRA AND THR. AND MUTAGENESIS OF LEU-891 AND
RP LEU-894.
RC TISSUE=Embryo;
RC MEDLINE=20148724; PubMed=10681503;
RA Cairn F., Antonson P., Pelto-Huikko M., Treurer E., Gustafson J.-A.;
RT "Cloning and characterization of RAP250, a nuclear receptor
RT coactivator."
RL J. Biol. Chem. 275:5308-5317(2000).
RN [4]
RP INTERACTION WITH RNP2.
RP MEDLINE=21638469; PubMed=11704680;
RA Jung D.-J., Na S.-Y., Na D.S., Lee J.W.;
RT "Molecular cloning and characterization of CAPER, a novel coactivator
RT of activating protein-1 and estrogen receptors."
RL J. Biol. Chem. 277:1229-1234(2002).
RC -FUNCTION: Nuclear receptor coactivator that directly binds nuclear
CC receptors and stimulates the transcriptional activities in a
CC hormone-dependent fashion. Coactivates expression in an agonist-
CC different nuclear receptors, such as for steroids (GR and ERs),
CC retinoids (RARs and RXRs), thyroid hormone (TRs), vitamin D3 (VDR)
CC and prostanoids (PPARs). Probably functions as a general
CC coactivator, rather than just a nuclear receptor coactivator. May
CC also be involved in the coactivation of the NF-kappa-B pathway.
CC May coactivate expression via a remodeling of chromatin and its
CC interaction with histone acetyltransferase proteins. Involved in
CC Placental, cardiac, hepatic and embryonic development.
CC -SUBUNIT: Monomer and homodimer. Interacts in vitro with the basal
CC transcription factors GTF2A and TBP, suggesting an autonomous
CC transactivation function. Interacts with NCOA1, CREB3, RBM14, the
CC histone acetyltransferase proteins EP300 and CREBBP, and with
CC methyltransferase proteins NCOA6 and HRMTL1 (By similarity).
CC Interacts with RNP2. Belongs to the ASC-2/NCOA6 complex (ASCOM),
CC which contains ASC-2/NCOA6, the retinoblastoma-binding protein
CC RBQ-3/RBBP5, alpha- and beta-tubulins, the trithorax group
CC proteins ML12 and ML13, and ASH2/ASCL2 (By similarity).
CC -SUBCELLULAR LOCATION: Nuclear.
CC -ALTERNATIVE PRODUCTS: Nuclear.
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=Q9JLI9-1; Sequence=Displayed;
CC Name=2;

```



```

Db 2834 IEDEKPVDLTLAGRRVACCDWYKLPFGSRCTAOP-----PTTLPEDFRGYDDHYGD 2888
Qy 1780 GSGSPHSFOLSK-SLPSONLSOS-----LSNFSNSVMSNDSEDEDLKLELR 1832
Db 2889 RSGSYGYGMGMKMSMDTLNSELGLFAYKSKSFDVQVGTDAVLTSGRVT----- 2943
Qy 1833 LRDRLKLTLOLQSFQKEHIESLYKLGKVPVAVIIPPAALSGRRRPTKSKGSRS 1892
Db 2944 -----SGEWMVYSSK-----TTGPYPETRQVIGIGI--STQYSGAR-----MVSSLS 2985
Qy 1893 SSLGKSPQLSGNSGSAASVHLHQQLH-----PQNPISGQNL--LQPLKPSPSD 1946
Db 2986 SPFGAGSVLRSSN--GVVYSSVATPIPTFAITTPGSIPTVYADLPLOTIDSVPLS 3043
Qy 1947 NL-----YSAFTS-----DGAISVPSLSAPGQGTSTVGAATN 1981
Db 3044 TLQONQPLPREYSFLTTMAAKKASTIIDITGTPPLTLESIAIPMLPATAS 3098

RESULT 14
M4K4 HUMAN
ID M4K4 HUMAN STANDARD; PRT; 1239 AA.
AC 095819; 075172; 09N87;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Mitogen-activated protein kinase kinase kinase 4 (EC 2.7.1.37)
DE (MAPK/ERK kinase kinase 4) (MEK kinase 4) (MEKK 4)
DE (HRK/CKC-like kinase HKK) (Nck interacting kinase).
GN MAP4K4 OR HKK OR NIK OR KTA0687.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OC NCBI_TaxID=9606;
OX [1]
RP SEQUENCE FROM N.A. (ISOFORMS 4 AND 5), FUNCTION, AND TISSUE
RP SPECIFICITY.
RC TISSUE=Macrophage;
RC MEDLINE=9107863; PubMed=9890973;
RX Yao Z., Zhou G., Wang X.S., Brown A., Diener K., Gan H., Tan T.-H.;
RT "A novel human STE20-related protein kinase, HGK, that specifically
RT activates the c-Jun N-terminal kinase signaling pathway.",
RT J. Biol. Chem. 274.2118-2125(1999).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Glioblastoma;
RC MEDLINE=22499944; PubMed=12612079;
RX Wright J.H., Wang X., Manning G., Lamere B.J., Le P., Zhu S.,
RA Khatri D., Flanagan P.M., Buckley S.D., Whyte D.B., Howlett A.R.,
RA Bischoff J.R., Lipson K.B., Jallat B.;
RT "The STE20 kinase HGK is broadly expressed in human tumor cells and
RT can modulate cellular transformation, invasion, and adhesion.",
RT Mol. Cell. Biol. 23.2068-2082(2003).
RN [3]
RP SEQUENCE OF 38-1239 FROM N.A. (ISOFORM 2).
RC TISSUE=Brain;
RC MEDLINE=98403880; PubMed=9734811;
RX Ishikawa K.-I., Nagase T., Suyama M., Miyajima N., Tanaka A.,
RA Kotani H., Nomura N., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. X.
RT The complete sequences of 100 new cDNA clones from brain which can
RT code for large proteins in vitro.",
RT DNA Res. 5:169-176(1998).
RN [4]
RP SEQUENCE OF 262-1239 FROM N.A. (ISOFORM 3).
RC TISSUE=Testis;
RC Duesterhoeft A., lauber J., Mewes H.-W., Gassenhuber J., Wiemann S.;
RA Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE OF 700-1239 FROM N.A.
RC TISSUE=Brain;
RC Saito T., Seki N., Hori T.;
RT "Isolation, expression profile and chromosome assignment of a novel

```

```

RT serine/threonine kinase gene.",
RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.
RN [6]
RP INTERACTION WITH SPG3A.
RX MEDLINE=22275851; PubMed=12387898;
RA Iuan Z., Zhang Y., Liu A., Man Y., Cheng L., Hu G.;
RT "A novel GTP-binding protein HGBP3 interacts with NIK/HGK.",
RL FEBS Lett. 530:233-238(2002).
CC -1- FUNCTION: Serine/threonine kinase that may play a role in the
CC response to environmental stress and cytokines such as TNF-alpha.
CC Appears to act upstream of the c-Jun N-terminal pathway.
CC -1- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC -1- COFACTOR: Magnesium.
CC -1- SUBUNIT: Interacts with the SH3 domain of the adapter proteins Nck
CC (by similarity). Binds, via its CNH regulatory domain, to the N-
CC terminal region of SPG3A.
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=5;
CC Name=1; Synonyms=Tumor-associated;
CC IsoId=O95819-1; Sequence=Displayed;
CC Name=2;
CC IsoId=O95819-2; Sequence=VSP_007054, VSP_007057;
CC Name=3;
CC IsoId=O95819-3; Sequence=VSP_007056, VSP_007057;
CC Name=4; Synonyms=HGK-S;
CC IsoId=O95819-4; Sequence=VSP_007054, VSP_007055, VSP_007057,
CC VSP_007058;
CC Name=5; Synonyms=HGK-L;
CC IsoId=O95819-5; Sequence=VSP_007054, VSP_007055, VSP_007056,
CC VSP_007057.
CC -1- TISSUE SPECIFICITY: Appears to be ubiquitous, expressed in all
CC tissue types examined. Isoform 5 appears to be more abundant in
CC the brain, isoform 4 is predominant in the liver, skeletal muscle
CC and placenta.
CC -1- SIMILARITY: Belongs to the Ser/Thr family of protein kinases.
CC STE20 subfamily.
CC -1- SIMILARITY: Contains 1 CNH domain.
CC -1- CAUTION: Ref.4 sequence differs from that shown due to a
CC frameshift in position 1151.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL; AF096300; ABD16137.1; -
CC EMBL; AY212247; AAO32626.1; -
CC EMBL; AB014587; BAA31662.1; -
CC EMBL; AL137755; CAB70907.1; ALT_FRAME.
CC EMBL; AB013385; BAA33714.1; -
CC Genew; HGNC:6866; MAP4K4.
CC MIM; 604665; -
CC GO; GO:0005524; F-ATP binding; IDA.
CC GO; GO:0004674; F-protein serine/threonine kinase activity; IDA.
CC GO; GO:0006468; F-protein amino acid phosphorylation; IDA.
CC GO; GO:0007243; F-protein kinase cascade; IDA.
CC GO; GO:0006950; P:response to stress; IDA.
CC InterPro; IPR001180; CItiron.
CC InterPro; IPR000719; Prot_kinase.
CC InterPro; IPR008271; Ser_thr_kin_AS.
CC InterPro; IPR002290; Ser_thr_kinase.
CC InterPro; IPR001245; Tyr_pkinase.
CC Pfam; PF00780; CNH; 1.
CC Pfam; PF00069; Pkinase; 1.
CC ProDom; PD000601; Prot_kinase; 1.
CC SMART; SM00036; CNH; 1.
CC SMART; SM00220; S_TKc; 1.
CC SMART; SM00219; TyKc; 1.
CC PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
CC PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.

```


RU J. Biol. Chem. 277:50497-50502(2002).
 CC -1- FUNCTION: May act as a scaffolding protein involved in the
 CC organization of synaptic active zones and in synaptic vesicle
 CC trafficking.
 CC -1- SUBUNIT: Interacts with Rabac1/Pral, RIMS2 and profilin.
 CC -1- SUBCELLULAR LOCATION: Concentrated at the presynaptic side of
 CC synaptic junctions.
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=1;
 CC IsoId=Q9QYX7-1; Sequence=Displayed;
 CC Name=2;
 CC IsoId=Q9QYX7-2; Sequence=VSP_003928, VSP_003929;
 CC -1- TISSUE SPECIFICITY: Highly expressed in brain. Low levels found in
 CC stomach. Not detected in other tissues analyzed including adrenal
 CC gland, testis and pancreas.
 CC -1- DOMAIN: C2 domain 1 is involved in binding calcium and
 CC phospholipids. Calcium binds with low affinity but with high
 CC specificity and induces a large conformational change.
 CC -1- SIMILARITY: Contains 2 C2 domains.
 CC -1- SIMILARITY: Contains 1 PDZ/DHR domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: Y19185; CAB60731.2; -;
 DR EMBL: Y19186; CAB60732.2; -;
 DR EMBL: AF181269; AA055786.2; -;
 DR HSSP: P04410; 1A25.
 DR MGD: MGI:1349390; Polo.
 DR GO: GO:0045202; C:synaptic junction; IDA.
 DR GO: GO:0005509; F:calcium ion binding; ISS.
 DR GO: GO:0005544; F:calcium-dependent phospholipid binding; ISS.
 DR GO: GO:0005522; F:profilin binding; IDA.
 DR GO: GO:0019933; P:CAMP-mediated signaling; IDA.
 DR GO: GO:0007010; P:cytoskeleton organization and biogenesis; IDA.
 DR GO: GO:0030073; P:insulin secretion; IDA.
 DR GO: GO:0017157; P:regulation of exocytosis; IDA.
 DR GO: GO:0016080; P:synaptic vesicle targeting; NAS.
 DR InterPro: IPR000008; C2.
 DR InterPro: IPR001478; PDZ.
 DR InterPro: IPR008899; Znf_piccolo.
 DR Pfam: PF00168; C2; 2.
 DR Pfam: PF00595; PDZ; 1.
 DR Pfam: PF05715; Zf_piccolo; 2.
 DR SMART: SM00239; C2; 2.
 DR SMART: SM00239; C2; 2.
 DR PROSITE: PS00499; C2 DOMAIN 1; 1.
 DR PROSITE: PS00004; C2 DOMAIN 2; 2.
 DR PROSITE: PS01006; PDZ; 1.
 KW Calcium/phospholipid-binding; Metal-binding; Zinc; Zinc-finger;
 KW Repeat; Alternative splicing.
 FT DOMAIN 371 470
 FT ZN FING 502 526
 FT ZN FING 967 990
 FT DOMAIN 2305 2329
 FT DOMAIN 4394 4488
 FT DOMAIN 4607 4705
 FT DOMAIN 4922 5012
 FT VARSPPLIC 4829 4833
 FT VARSPPLIC 4834 5038
 FT SEQUENCE 5038 AA; 547600 MM; DADA460CF3B40888 CRC64;
 Query Match 3.7%; Score 404.5; DB 1; Length 5038;
 Best Local Similarity 19.4%; Pred. No. 2.6e-06;

Matches 399; Conservative 273; Mismatches 737; Indels 653; Gaps 91;
 QY 8 KQSTPGSLPLSPAPAPKNGSSSDSVGEKLGAAADAVTGRTEVRRRTMD-KQSR 66
 DB 1041 KASVP-----APAPPPQKPTLAHAKGK-----KETEVAKETKQPEKETSEIKTP 1091
 QY 67 GAATTTTHERFRFRSVYC-----DSNATLELPGPLSLPPPSI 107
 DB 1092 PAVATDQKLESEVTKLVSVLPEKKPEEEKALPADKKEKPPAPAPAPLEKKPIDDD 1151
 QY 108 PAAPVQAPPEPRHRE-TVATATSOVQAPP-----AAAPGRPA-----VAGPAPDS 154
 DB 1152 QKLPPDAKPSAEEBEKDLKAHVQIPESGPICKVSLAEQEQPPTREDLPGAPRP 1211
 QY 155 TVPSTSKD-RPVQSPSLVSGKEEPPAPASGGGSAKEPQERSQODDIELETKA-- 211
 DB 1212 TLPRDQKESRDVYQPOAEGT-----AKEERG-----EPKDKTEKEDSDTSSSQP 1260
 QY 212 -----VQMSNDRFLKPDVLEIGRGSFKTVYKGLDTETTVAVAMELODRKLTSEKQ 263
 DB 1261 KSPQGLSDTGYSDG-----ISGLGEPISLP----- 1288
 QY 264 RFKEEAMKLGLOHPNIVRFYDSWEST-----VKGKCCIVVTEMTSGTKYIKRF 316
 DB 1289 --SEKDLKGLKQKDSF-----SQESSPPSSDLAKLESTVLEIAQAS-TLVGERAKP 1340
 QY 317 KWKIKYLRWCRQLIKGLQFLHTRTPPIIHRDLK-CDNIFITGPTGSVK----- 365
 DB 1341 KTPQKVPSPQPODQKQTPSETRDISISBEELIKESQEKVIVSKDSAGCFPRKHKE 1400
 QY 366 -----IGDLGLATLKRAFSFASVIGTPFPMAPPEWEEKYDSDVYAFGMCLMATSEYP 421
 DB 1401 NPELVLDL--SPRRASY-DSVEDSSESENSPARRRKRTSIG-----SSSSEYK 1447
 QY 422 YSECONAAQ-----IYRRV-----TSGVKPSFDVVALPEKKEIEGCIRO----- 462
 DB 1448 QEISQSGSEBEDIIRKQIIEMSADEDSGEDEFIRSQKELIGVYESQKREBTYKGGK 1507
 QY 463 ---NKERYSIKOLNHAFPOEETGVRLAEEDDEKIAKILRIEDIKLKGKYN 519
 DB 1508 SPAGKRRLRLTK--SSTSPDDAGRRHSHWDEDD----- 1539
 QY 520 EALPEFPDLE-RDVPDVAQEWESGYVCEGHKT--MAKALDRVSLIRKKEQOLVR 576
 DB 1540 ETFDESPELKFRYTKSQESELLELVAGGGGLRRKKTIELNSVTDKYS--AESQKKTLL 1596
 QY 577 EDEKKKEESSILKQOVQSSASQGTGKQLPAS--TGIPLTASTASVSTQVEPREPEPA 634
 DB 1597 YFDEBELEMEGLTDSPEDRSGR-GSSSLHASSFTPTGTSVSSLDSDSSPSHKKG 1655
 QY 635 DHOQLOIQQPSISVLSGTVDSGGSSVFTSRVSSQQTVSQSGHQEQAHSQTGVPGH 694
 DB 1656 ESKQQRKARHSHGFLPTIEDSEEEELREBEELKEQEKRELEQQQRKSSSKSKXD 1715
 QY 695 PSTVQAQ-SQPHGVPPSSVNAQSGQSGSSSLGVSSQPIQHPDQOQGIQQTAPPOQ 753
 DB 1716 KQELRAQRREPRKTPPSNL-----PIEDASPEEELRQALEEME 1755
 QY 754 TVQVSLQSTSTSEATTAPVQSPQAPVLPQVNSAGKSTQGVQVAPAEVAVVQPPAT 813
 DB 1756 LHRSSCSESPSEI--SDPEGFISPEKITIVQKYKLPVLSVLSPTDEDSVWQKEA 1812
 QY 814 QPTTASSVDSAHSDVAGSGMDGNNENPSSSGRHGRRTTKRHRYKSVRSRHEKTSRPK 873
 DB 1813 Q-----KALKSAEMEYEMMHKPI-----K 1832
 QY 874 LRLINVSXKGRVVECO-----LETNRKKV--TKFKLDGMPPEIATIVNNDPI 923
 DB 1833 YKAFPAANERDEVEFEKPLYGGMLEDIYISLVDYTVAGSDG-----SLATROD-- 1883
 QY 924 LAIERESFVDQ-----VREIIEKADEMISEDVSVPEDQGIESTQCKDDYFSGSGQ 975
 DB 1884 ---EONGMMPQGRGROKIRLRQI--YDDPWQKTTDQKEFYE-LESLH----- 1926

Blank sheet

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 24, 2004, 00:59:43 ; Search time 195 Seconds
(without alignments)
3456.138 Million cell updates/sec

Title: US-10-010-720-14

Sequence: 1 MEGGAALKKSTPGSLFLSP.....NISLQKSIKNPPGSLNRTT 2136

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : Listing first 45 summaries

1: SP_TREMBL_25:*
2: sp_archaea:*
3: sp_bacteria:*
4: sp_fungi:*
5: sp_human:*
6: sp_invertebrate:*
7: sp_mammal:*
8: sp_mhc:*
9: sp_organelle:*
10: sp_phage:*
11: sp_plant:*
12: sp_rodent:*
13: sp_virus:*
14: sp_vertebrate:*
15: sp_unclassified:*
16: sp_virus:*
17: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	10677	98.8	2382	4	Q9H4A3
2	10435	96.5	2066	4	O15052
3	9204	85.1	2126	11	Q9JH7
4	3408	31.5	670	4	Q9JH7
5	2630	24.3	2296	4	O9Y3S1
6	2615	24.2	2219	4	O9C0A3
7	2463.5	22.6	1800	4	O8TCX6
8	2439	22.6	1743	4	O9BYP7
9	2126	19.7	1243	4	O96J92
10	2083	19.3	1222	11	O810H4
11	2082	19.3	1210	11	O811R5
12	2068	19.1	1222	11	O80UR6
13	1974	18.3	393	4	O8N673
14	1911.5	17.7	1048	11	O80XN2
15	1890.5	17.5	1552	11	O7TPK6
16	1804.5	16.7	779	4	Q9H3P4

17	1682.5	15.6	417	11	O8K2Q3	O8K2q3 mus musculus
18	1655.5	15.3	1839	5	O18657	O18657 caenorhabditis
19	1610	14.9	1678	5	O81127	O81127 caenorhabditis
20	1507.5	13.9	954	11	O80XB5	O80XB5 mus musculus
21	1426	13.2	1557	11	O811F2	O811f2 mus musculus
22	1362.5	12.6	663	4	O8N8X3	O8N8X3 homo sapien
23	1238	11.5	2308	5	O9YPI7	O9YPI7 drosophila
24	1030.5	9.5	1213	4	O9HCX6	O9HCX6 homo sapien
25	945	8.7	212	4	O9BYSS	O9BYSS homo sapien
26	915	8.5	173	4	O96CZ6	O96CZ6 homo sapien
27	856	7.9	176	11	O7TPR2	O7TPR2 mus musculus
28	768	7.1	591	10	O84R51	O84R51 medicago sa
29	761.5	7.0	563	10	O944Q0	O944Q0 arabidopsis
30	761	7.0	608	10	O9FHT4	O9FHT4 arabidopsis
31	759.5	7.0	550	10	O8S8Y7	O8S8Y7 arabidopsis
32	755.5	7.0	557	10	O8LST2	O8LST2 arabidopsis
33	755.5	7.0	666	10	O9PDV6	O9PDV6 fagus sylvia
34	752.5	7.0	567	10	O8S8Y8	O8S8Y8 arabidopsis
35	750.5	6.9	578	10	O9LUV8	O9LUV8 arabidopsis
36	746	6.9	524	10	O8RXE5	O8RXE5 arabidopsis
37	745.5	6.9	516	10	O9STK6	O9STK6 arabidopsis
38	742.5	6.9	568	10	O8S8Y9	O8S8Y9 arabidopsis
39	740.5	6.8	571	10	O9LSA2	O9LSA2 arabidopsis
40	715.5	6.6	539	10	O9CSN1	O9CSN1 arabidopsis
41	709.5	6.6	542	10	O9M9B6	O9M9B6 arabidopsis
42	702.5	6.5	542	10	O94HK9	O94HK9 oryza sativ
43	702.5	6.5	542	10	O7XB94	O7XB94 oryza sativ
44	702.5	6.5	547	10	O9SCU5	O9SCU5 arabidopsis
45	702.5	6.5	549	10	O9XF55	O9XF55 arabidopsis

ALIGNMENTS

RESULT 1

Q9H4A3

PRELIMINARY;

PRT; 2382 AA.

ID Q9H4A3
AC Q9H4A3;
DT 01-MAR-2001 (TREMURel. 16, Created)
DT 01-MAR-2001 (TREMURel. 16, Last sequence update)
DT 01-OCR-2003 (TREMURel. 25, Last annotation update)
DE Hypothetical protein.
GN WNK1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Heart;
RA Verisimo F., Jordan P.;
RT "A novel human protein kinase gene, WNK1: cDNA cloning, expression and genomic structure."
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; AJ296290; CAC15059.1;
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR00719; Prot. kinase.
DR InterPro; IPR008271; Ser_thr_kin_AS.
DR Pfam; PF00069; Pkinase; 1.
DR Prodom; PD00001; Prot. kinase; 1.
DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS0108; PROTEIN_KINASE_ST; 1.
KW Hypothetical protein; ATP-binding; transferase.
KW Serine/threonine-protein kinase; transferase.
SQ SEQUENCE 2382 AA; 250755 MW; ADDA6C8554B7CA CRC64;

Query Match 98.8%; Score 10677; DB 4; Length 2382;
Best Local Similarity 89.6%; Pred. No. 0;
Matches 2135; Conservative 0; Mismatches 1; Indels 246; Gaps 1;

QY 1 MSGAAEKOSSTPGSLFLSPAPAPKNGSSSDSVGEKLGAAAADAVTGRTEYRRRTT 60
Db 1 MSGAAEKOSSTPGSLFLSPAPAPKNGSSSDSVGEKLGAAAADAVTGRTEYRRRTT 60
QY 61 MOKDSRGAATTTTTEHRRFRFRSVICDSNAKLELPLSLPSPSIPAAVQSAPEBH 120
Db 61 MOKDSRGAATTTTTEHRRFRFRSVICDSNAKLELPLSLPSPSIPAAVQSAPEBH 120
QY 121 REETATATSOVAOOPRAAAGPAAVGPASTVPSTSKDRVSOPLVSGSEEP 180
Db 121 REETATATSOVAOOPRAAAGPAAVGPASTVPSTSKDRVSOPLVSGSEEP 180
QY 181 ARSGGGGSAKEPOEERSQOQDDIELEETKAVGMSNDGRFLKFDIEIGRSEFKTYKGLD 240
Db 181 ARSGGGGSAKEPOEERSQOQDDIELEETKAVGMSNDGRFLKFDIEIGRSEFKTYKGLD 240
QY 241 TETTVAVMCELQORUKLTKSERORKEAEMLKGIQHNIVRFYDSWESTVYKXKCIYLV 300
Db 241 TETTVAVMCELQORUKLTKSERORKEAEMLKGIQHNIVRFYDSWESTVYKXKCIYLV 300
QY 301 TELMTSGTLTKYLRKFYMKIKVLRSWCROILKGLQFLHTPTPLIHRDLKCDNIPTGP 360
Db 301 TELMTSGTLTKYLRKFYMKIKVLRSWCROILKGLQFLHTPTPLIHRDLKCDNIPTGP 360
QY 361 TGSYKIGDLGLATLKRAFAKSVIGTPEFMAPEMYEERKYDESVDVYAFGCMLEMAISEY 420
Db 361 TGSYKIGDLGLATLKRAFAKSVIGTPEFMAPEMYEERKYDESVDVYAFGCMLEMAISEY 420
QY 421 PYSCQNAOQYRRVNTSGVKNASFDKVAIPVXKEIIEGCIKQNDERYSIKDLINHAFFQ 480
Db 421 PYSCQNAOQYRRVNTSGVKNASFDKVAIPVXKEIIEGCIKQNDERYSIKDLINHAFFQ 480
QY 481 BETGVARELAEDDEGEKIAIKLMLRIDIKLGKQYKQNDERIEISFLERVPEDVAQEM 540
Db 481 BETGVARELAEDDEGEKIAIKLMLRIDIKLGKQYKQNDERIEISFLERVPEDVAQEM 540
QY 541 VESGYVEGECHHTMAKAIKDRVSLIKRRRQOLVREOEKKQBEESLKOQVQSSASQ 600
Db 541 VESGYVEGECHHTMAKAIKDRVSLIKRRRQOLVREOEKKQBEESLKOQVQSSASQ 600
QY 601 TGIKOLPSASTGIPTAATTSASVSTOVEPEPEADHQOQLOQOQPSISVSDGTVDGSG 660
Db 601 TGIKOLPSASTGIPTAATTSASVSTOVEPEPEADHQOQLOQOQPSISVSDGTVDGSG 660
QY 661 SSVFTEBSRVSSQOTVSGSQHEQAHSITGYPGHPSTIVQASQPHGVYPPSSVAQGGSG 720
Db 661 SSVFTEBSRVSSQOTVSGSQHEQAHSITGYPGHPSTIVQASQPHGVYPPSSVAQGGSG 720
QY 721 QPSSSSSLTGVSSQPIQHPQOQOQGIQOTAPQOTVQYSLQSTSSSATTAAQVQSQPAR 780
Db 721 QPSSSSSLTGVSSQPIQHPQOQOQGIQOTAPQOTVQYSLQSTSSSATTAAQVQSQPAR 780
QY 781 QVLPQVSAKQOLVSOQVPPTIGQEPQIPVATOPSVPVHSGAFLPVGQPLPTPLPQYR 840
Db 781 QVLPQVSAKQOLVSOQVPPTIGQEPQIPVATOPSVPVHSGAFLPVGQPLPTPLPQYR 840
QY 841 VSOQIPISTPHVSTAOQTFSSSLPTITMAAGITQPLLTILASSATTALPGVSTVPSQLPTLL 900
Db 841 VSOQIPISTPHVSTAOQTFSSSLPTITMAAGITQPLLTILASSATTALPGVSTVPSQLPTLL 900
QY 901 QPVTOLPSQVHPQLOLPAVQSMGIPANLGQAAEVPPLSGDVLVYQGFPPRLPQYPGDSNI 960
Db 901 QPVTOLPSQVHPQLOLPAVQSMGIPANLGQAAEVPPLSGDVLVYQGFPPRLPQYPGDSNI 960
QY 961 AASSNVAVCIHSTVUSPPMPTVELATPGYFPTVQPYVESNLVPMGCVGQVQVQSPG 1020
Db 961 AASSNVAVCIHSTVUSPPMPTVELATPGYFPTVQPYVESNLVPMGCVGQVQVQSPG 1020
QY 1021 GSLAQAPTTSSQOAVLESTQGVSAVAPAPVAVQAPQATQPTTLASSVDSASDVASGMS 1080
Db 1021 GSLAQAPTTSSQOAVLESTQGVSAVAPAPVAVQAPQATQPTTLASSVDSASDVASGMS 1080

QY 835 DGENVVPSSSGHREGRITKRRHRSVRSRSRHEKTSRPRKLILVNSKNGRVVECOLETH 894
Db 1081 DGENVVPSSSGHREGRITKRRHRSVRSRSRHEKTSRPRKLILVNSKNGRVVECOLETH 1140
QY 895 NRKWTFFKEDLDGNDPEELATIMVNDPLATERSSFPVDQVREIEKADEMISEVSEVP 954
Db 1141 NRKWTFFKEDLDGNDPEELATIMVNDPLATERSSFPVDQVREIEKADEMISEVSEVP 1200
QY 955 EGDQLESLOQKDDYGFSGSQKLEGEFKQPIPASMPQOIGIPTSLQVYHSAGRFV 1014
Db 1201 EGDQLESLOQKDDYGFSGSQKLEGEFKQPIPASMPQOIGIPTSLQVYHSAGRFV 1260
QY 1015 SPVESRLRESKVPSEITTDVAASTASPGMNLSSHASLSLQAFSELRAQTEGPN 1074
Db 1261 SPVESRLRESKVPSEITTDVAASTASPGMNLSSHASLSLQAFSELRAQTEGPN 1320
QY 1075 TAPNFSGHTGTPFPVPPFLSSIGVPTTAATKVPVATSSPPNDISTSVQSEVTYTE 1134
Db 1321 TAPNFSGHTGTPFPVPPFLSSIGVPTTAATKVPVATSSPPNDISTSVQSEVTYTE 1380
QY 1135 EGIAGVATSTGVTSGGLPIPPVSESPVSSVSITIPAVVSISSPSLOVPTSEI 1194
Db 1381 EGIAGVATSTGVTSGGLPIPPVSESPVSSVSITIPAVVSISSPSLOVPTSEI 1440
QY 1195 VVSSTALYPSVTGATSAAGSTATTPGKPPAVVSOQAAGSTTVGATLISVSTTSPS 1254
Db 1441 VVSSTALYPSVTGATSAAGSTATTPGKPPAVVSOQAAGSTTVGATLISVSTTSPS 1500
QY 1255 TASQSLTOLSSSTPTPLAETVVVVSASHSDKTSSTTGAFSLASASSSSPBAQVSY 1314
Db 1501 TASQSLTOLSSSTPTPLAETVVVVSASHSDKTSSTTGAFSLASASSSSPBAQVSY 1560
QY 1315 ISOPGGLHPIVPSVASTPILPQAAGTSTPULLQVPSIPPLVQVANYANVQOQLIHS 1374
Db 1561 ISOPGGLHPIVPSVASTPILPQAAGTSTPULLQVPSIPPLVQVANYANVQOQLIHS 1620
QY 1375 QPQALLPQNPQHPHCEBVDSDTOPKAGIDIDITLEBKLSLSEHSSSGAQAHSVLET 1434
Db 1621 QPQALLPQNPQHPHCEBVDSDTOPKAGIDIDITLEBKLSLSEHSSSGAQAHSVLET 1680
QY 1435 SLVIESVTPEGIPPTVAVAPSKLSTSTGCLPPTNPLIGVVALPVYVYTPQGVSTPVT 1494
Db 1681 SLVIESVTPEGIPPTVAVAPSKLSTSTGCLPPTNPLIGVVALPVYVYTPQGVSTPVT 1740
QY 1495 TTSQVKGTAESPPLTKAPVLEVGTELPAGTLPSQOLPPFPSPSILQSQPLEDDAOL 1554
Db 1741 TTSQVKGTAESPPLTKAPVLEVGTELPAGTLPSQOLPPFPSPSILQSQPLEDDAOL 1800
QY 1555 RRTLSPEXITVTSAVGPVMAAPATATTEAGTOPQKVSOVKEGPVATSSGAGVFKXGRF 1614
Db 1801 RRTLSPEXITVTSAVGPVMAAPATATTEAGTOPQKVSOVKEGPVATSSGAGVFKXGRF 1860
QY 1615 QVSVADGAQKQKNSSEDAKSVHFESSTSESSVSSSPESTLVPEPVGITIPGISSD 1674
Db 1861 QVSVADGAQKQKNSSEDAKSVHFESSTSESSVSSSPESTLVPEPVGITIPGISSD 1920
QY 1675 VPESAKHTTASBAKSDTGPQTVKGRFOVTTANKVGRFVSXKTEDKITDTYKKGVPYASP 1734
Db 1921 VPESAKHTTASBAKSDTGPQTVKGRFOVTTANKVGRFVSXKTEDKITDTYKKGVPYASP 1980
QY 1735 FMDLEQAVLPAVIPKKEKPELSEPSHLNGSSDPKPAFLSRVDDGSGSPHSHQJSSGS 1794
Db 1981 FMDLEQAVLPAVIPKKEKPELSEPSHLNGSSDPKPAFLSRVDDGSGSPHSHQJSSGS 2040
QY 1795 LPSQNLSSLSNSFNSSVMSDNESDIEDEDKLRLRDLKHEIDLOJOSROKHETES 1854
Db 2041 LPSQNLSSLSNSFNSSVMSDNESDIEDEDKLRLRDLKHEIDLOJOSROKHETES 2100
QY 1855 LYTXYKGPVPAVITPPAAPLSGRRRRPTKSKSKSSRSSSLGNKSPOLSGNLGGQSAASV 1914
Db 2101 LYTXYKGPVPAVITPPAAPLSGRRRRPTKSKSKSSRSSSLGNKSPOLSGNLGGQSAASV 2160
QY 1915 LHPQOTLHPPGNIPESQONQLOPLKPSBSSDNLXSAFTSDGALVPSLSAPGQGTSTTN 1974

Dc		2161	LHPQQLTLPKPNIPESGQNQDLPLPKRSPSSDNLVYSAFISDGAI SVPSLSAPeQGTSSTN	2220
Oy		1975	TVGATVNSQAQAOPAMPMTSSRKGTFFDDHLKLVDMNARDAMNLGSRGSGKHMYEGRG	2034
Dd		2221	TVGATVNSQAQAOPAMPMTSSRKCTFTDDLHKLVDMNARAMNLGSRGSGKHMYEGRG	2280
Oy		2035	MARRSAPGOLCISSMTSNLGGSAPIISAASATSIGHPTYSKCPOQOYGFPATPGAGMSGT	2094
Dd		2281	MARRSAPGOLCISMTSNLGGSAPIISAASATSIGHPTYSKCPCPOQYGFPA TPGAOMSGT	2340
Oy		2095	GGPAPOLIGOFQPVGTASLQNFNISNQKSI SNRGSNIETT	2136
Dd		2341	GGPAPOLIGOFQPVGTASLQNFNISNLQKSINPDSGNLRTT	2382
RESULT 2				
ID	015052		PRELIMINARY; PRT, 2066 AA.	
AC	015052;			
DT	01-JAN-1998	(TREMBLrel. 05, Created)		
DT	01-OCT-2002	(TREMBLrel. 22, Last sequence update)		
DT	01-OCT-2003	(TREMBLrel. 25, Last annotation update)		
DE	Hypothetical protein KIAA0344 (Fragment).	KIAA0344.		
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.			
OX	NCB1_Taxid=9606;			
RN	(1)			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Brain;			
RA	MEDLINE=97449984; PubMed=9205841;			
RA	Nagase T., Ishikawa K., Nakajima D., Ohira M., Seki N., Miyajima N.,			
RA	Tanaka A., Kotani H., Nomura N., Ohara O.;			
RT	"Prediction of the coding sequences of unidentified human genes. VII.			
RT	The complete sequences of 100 new cDNA clones from brain which can			
RL	code for large proteins in vitro."			
RL	DNA Res. 4:141-150(1997).			
DR	EMBL; AB002342; BAA20802.2; -			
DR	Genev; HGNC:14540; PRKNMKI.			
DR	GO; GO:0005524; F:ATP binding; IEA.			
DR	GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.			
DR	GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.			
DR	GO; GO:0016740; F:transferase activity; IEA.			
DR	GO; GO:0004668; P:protein amino acid phosphorylation; IEA.			
DR	InterPro; IPR000719; Prot_kinase.			
DR	InterPro; IPR002290; Ser_thr_kinase.			
DR	InterPro; IPR008271; Ser_thr_pkin_AS.			
DR	InterPro; IPR001245; Tyr_pkinase.			
DR	Pfam; PF00069; pkinase; 1.			
DR	Prodom; PD000001; Prot_kinase; 1.			
DR	SMART; SM00220; S_TKc; 1.			
DR	SMART; SM00219; TYKc; 1.			
DR	PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.			
DR	PROSITE; PSS0108; PROTEIN_KINASE_ST; 1.			
KM	Hypotheical protein; ATP-binding; Transferase.			
FT	NON TER			
SQ	SEQUENCE	2066 AA; 218551 MW; EC59F36363d45DU3 CRC64;		
Query Match				
	Best Local Similarity	96.5%; Score 10435; DB 4; Length 2066;		
	Matches 2065; Conservativity	0; Mismatches 1; Indels 2; Gaps 2.		
Oy		69 AATTITTHRRFFRRSYICDSNATALEPLGLPISTLOPSTIPAANVPQAPPEPRREETVAT	128	
Dd		1 AATTITTHRRFFRRSYICDSNATALEPLGLPISTLOPSTIPAANVPQAPPEPRREETVAT	60	
Oy		129 ATSOVAQQPPAALAAPGEAVAGCAPSVTSSTSKDPVSQPSTLYSGKEEPPPARSGGGG	188	
Dd		61 ATSOVAQQPPAALAAPGEAVAGCAPSVTSSTSKDPVSQPSTLYSGKEEPPPARSGGGG	120	
Oy		189 SAKEPSEERSQQDDIEBELETFAVGMNDGRFLKFIDEIKGRGSFKTYIVAGLDTETTVEYA	248	

D	b	121	SAPEPBEHSQOQDDIEELJETAVAGNSNDRFLKEDIEIGRGSFXTVVXGIDTETTTVEA	180
Q	y	249	WCELODRKLTKSRQRFPEAEMLKELQHPNIYRFPDSESVTKKKCCIVLATELMTSGT	308
D	b	181	WCELODRKLTKSRQRFPEAEMLKELQHPNIYRFPDSESVTKKKCCIVLATELMTSGT	240
Q	y	309	LKTYLKFEXMKIKVLRSMCROTLKQLEFLHRTPEPIHRDLKCNIFITGTGSAVKIGD	368
D	b	241	LKTYLKFEXMKIKVLRSMCROTLKQLEFLHRTPEPIHRDLKCNIFITGTGSAVKIGD	300
Q	y	369	LGIALTLKRASFASVIGTPEFMAPEMYEEKYDESDVYAFQMCOTEMATSEYFSECONA	428
D	b	301	LGIALTLKRASFASVIGTPEFMAPEMYEEKYDESDVYAFQMCOTEMATSEYFSECONA	360
Q	y	429	AQIYRRTSGVKASPPKVAIPVKKIINGCIRONDERYSIKDLINHPFOEBGVRYE	488
D	b	361	AQIYRRTSGVKASPPKVAIPVKKIINGCIRONDERYSIKDLINHPFOEBGVRYE	420
Q	y	489	LAEDDEKIALKMLTIEDIKLKKGKYNKEAIEFSFLDERVPEDVAQEWMSGYCE	548
D	b	421	LAEDDEKIALKMLTIEDIKLKKGKYNKEAIEFSFLDERVPEDVAQEWMSGYCE	480
Q	y	549	GDHKTMAKAIKDRVSLIKRYREPOLYREBEKKQPESSLKQOVQSSASQTKQOLPS	608
D	b	481	GDHKTMAKAIKDRVSLIKRYREPOLYREBEKKQPESSLKQOVQSSASQTKQOLPS	540
Q	y	609	ASTGIPATSTASVSQVEPEPEADQHOLOQOQPSISVLSDGCVNDSQGSSTVTEER	668
D	b	541	ASTGIPATSTASVSQVEPEPEADQHOLOQOQPSISVLSDGCVNDSQGSSTVTEER	600
Q	y	669	VSSQQTASVYSGQHEQAHSTGTPGHIPSTVQAOQSOPHGVYPRSSVAQGSQGPSSSLT	728
D	b	601	VSSQQTASVYSGQHEQAHSTGTPGHIPSTVQAOQSOPHGVYPRSSVAQGSQGPSSSLT	660
Q	y	729	GVSSSQPIQHQOQOQGIQOTAPPCQTVQYSLSQSTSSSEKTAQOPVSOPAPQVYLQVSA	788
D	b	661	GVSSSQPIQHQOQOQGIQOTAPPCQTVQYSLSQSTSSSEKTAQOPVSOPAPQVYLQVSA	719
Q	y	789	GKOSQVQSVQVAPAEPAVVAPOCATQCTTLASVDSASHSDVASMDSGNENVBSSGRHE	848
D	b	720	GKOSQVQSVQVAPAEPAVVAPOCATQCTTLASVDSASHSDVASMDSGNENVBSSGRHE	779
Q	y	849	GTTKRRHRSKVRBSRHEKTSRPKLKIILVNSKGBDVECOLETHNRKNVTKFPLDGD	908
D	b	780	GTTKRRHRSKVRBSRHEKTSRPKLKIILVNSKGBDVECOLETHNRKNVTKFPLDGD	839
Q	y	909	NBEELATIMANNDFIATERESFVDQREITEKDEMISEVSVSEPEGDGLESLOKOD	968
D	b	840	NBEELATIMANNDFIATERESFVDQREITEKDEMISEVSVSEPEGDGLESLOKOD	899
Q	y	969	YGFSGSQKLEGFHFKOPIBASSMPOQIGITSSLTQVYHSAGRFIVPVBSRLREKVF	1028
D	b	900	YGFSGSQKLEGFHFKOPIBASSMPOQIGITSSLTQVYHSAGRFIVPVBSRLREKVF	959
Q	y	1029	PBEITDYAASSTAOSSPGMNLSHSASSLSLOAFSELRAQMTBEPNTPAPVFSHTGTFP	1088
D	b	960	PBEITDYAASSTAOSSPGMNLSHSASSLSLOAFSELRAQMTBEPNTPAPVFSHTGTFP	1019
Q	y	1089	VVPPEFISIAQVPTTAAATAPVPATSPSPNDISTSVIOSEVTBTEEGIAGVATSTGVT	1148
D	b	1020	VVPPEFISIAQVPTTAAATAPVPATSPSPNDISTSVIOSEVTBTEEGIAGVATSTGVT	1079
Q	y	1149	SGGLPIPVSESPVSVSSVSTTTPAVYSISTSSLOVPSTSEIIVSSALPASTVS	1208
D	b	1080	SGGLPIPVSESPVSVSSVSTTTPAVYSISTSSLOVPSTSEIIVSSALPASTVS	1139
Q	y	1209	ATSASASGSGTAPGKPPAVVVSQQAAGSTVGCATLTSVITSSPESTASOLUSSTSS	1268
D	b	1140	ATSASASGSGTAPGKPPAVVVSQQAAGSTVGCATLTSVITSSPESTASOLUSSTSS	1199
Q	y	1269	TPPLAETVVVSAHSLDKTSHSSTGLAFSLAPSSSSSPGACVSVYSIQPGGLHPVLPYS	1328

Db 1200 TPTLAETVAVVSAHSDKTSHTSSTGLAFLSAPSSSSSPGAGVSSYISQPGIHLPLVIPS 1259
QY 1329 VIATPILIPQAGPTSTPLPQVPSIPPLVQVAVVAVVAVQOQLHSOPQALLPNOPTH 1388
Db 1260 VIATPILIPQAGPTSTPLPQVPSIPPLVQVAVVAVVAVQOQLHSOPQALLPNOPTH 1319
QY 1389 CPVVSSTPQKAPGIDDKTLEKRLSLPSEHSSSGAQAHSVLETSVIESTVPGIPT 1448
Db 1320 CPVVSSTPQKAPGIDDKTLEKRLSLPSEHSSSGAQAHSVLETSVIESTVPGIPT 1379
QY 1449 TAVAPSKLTSTSTSCLPPTNPLGTVLPLVTVVTPPQVSTPSTTSVGVKPGTASBP 1508
Db 1380 TAVAPSKLTSTSTSCLPPTNPLGTVLPLVTVVTPPQVSTPSTTSVGVKPGTASBP 1439
QY 1509 PLTAPAVLPVGTETLPAAGTLPSBOLPFPFGPSLTQSOQLPDLIDQRLRTLSPEXITVISA 1568
Db 1440 PLTAPAVLPVGTETLPAAGTLPSBOLPFPFGPSLTQSOQLPDLIDQRLRTLSPEXITVISA 1499
QY 1569 VGPVSMAPTAITENGTPQKGVGVGVKGPVLTATSSGAVFRMGKRFQVSVADGAQKXGK 1628
Db 1500 VGPVSMAPTAITENGTPQKGVGVGVKGPVLTATSSGAVFRMGKRFQVSVADGAQKXGK 1558
QY 1629 NKSEDAKSVFEESTSESSSVSSSPSTLYKPPNGITITIGISDVPEAHKTTASBAK 1688
Db 1559 NKSEDAKSVFEESTSESSSVSSSPSTLYKPPNGITITIGISDVPEAHKTTASBAK 1618
QY 1689 SDTQOPTKVGKRFQVTTANKVGRFSVSKTEDKIDTKKKGCVASPPFMDLEQAVLPVAVIP 1748
Db 1619 SDTQOPTKVGKRFQVTTANKVGRFSVSKTEDKIDTKKKGCVASPPFMDLEQAVLPVAVIP 1678
QY 1749 KKEKPELSESHLNGPSSDPEAFLSRDVDGSGSPHSPHQLSKSLPSONLSQSLNSF 1808
Db 1679 KKEKPELSESHLNGPSSDPEAFLSRDVDGSGSPHSPHQLSKSLPSONLSQSLNSF 1738
QY 1809 NSSYMSDNESEDIEDDLKELRLRDKHLKEIODLSROKHESLTYTLGKVPVAVII 1868
Db 1739 NSSYMSDNESEDIEDDLKELRLRDKHLKEIODLSROKHESLTYTLGKVPVAVII 1798
QY 1869 PPAAPLSGRRRRPRTKSGKSSRSSSLGANKSPOLSGMLSGQSAASVHLPPQTLHPNINIP 1928
Db 1799 PPAAPLSGRRRRPRTKSGKSSRSSSLGANKSPOLSGMLSGQSAASVHLPPQTLHPNINIP 1858
QY 1929 ESGONOLLOPLKSPSSDNLKYSAPTSIGALISVPSLAPGQGTSSWTVGATVNSQAQAQ 1988
Db 1859 ESGONOLLOPLKSPSSDNLKYSAPTSIGALISVPSLAPGQGTSSWTVGATVNSQAQAQ 1918
QY 1989 PPAATSSRKGTFTDDHLKLVNNAARDAMNLGRRGSGHNNYEGPGMAKRFSAPOGLCTIS 2048
Db 1919 PPAATSSRKGTFTDDHLKLVNNAARDAMNLGRRGSGHNNYEGPGMAKRFSAPOGLCTIS 1978
QY 2049 MTSNLGSAPIASAATSLGHTFTKSCPPQOYGPATPFGAOWSGTGPAQPLGQFOFV 2108
Db 1979 MTSNLGSAPIASAATSLGHTFTKSCPPQOYGPATPFGAOWSGTGPAQPLGQFOFV 2038
QY 2109 GTASLONFNINSLOKGISNPPGSLNRTT 2136
Db 2039 GTASLONFNINSLOKGISNPPGSLNRTT 2066

RP SEQUENCE FROM N.A.
RX MEDLINE=2028750; PubMed=10828064;
RA Xu B., English J.M., Wilsbacher J.L., Stippec S., Goldsmith E.J.,
RT Cobb M.H.;
RT "WNKI, a Novel Mammalian Serine/Threonine Protein Kinase Lacking the
RT Catalytic Lysine in Subdomain II";
RL J. Biol. Chem. 275:16795-16801(2000).
CC -I- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; AF227741; AAF74258.1; -;
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR000719; Prot. Kinase.
DR InterPro; IPR008271; Ser. Thr. Kin. AS.
DR Pfam; PF00069; Pkinase; 1.
DR ProDom; PD000001; Prot. kinase; 1.
DR PROSITE; PS00011; PROTEIN KINASE DOM.; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST.; 1.
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 2126 AA; 225216 MW; 7137BC45A54402F CRC64;

Query Match 85.1%; Score 9204; DB 11; Length 2126;
Best Local Similarity 85.6%; Pred. No. 0;
Matches 1838; Conservative 83; Mismatches 193; Indels 34; Gaps 15;

QY 1 MSGGAARKSGSTPGLSLSPAPAPKNGSSSDSSVGBKLGAADAVTGRTEYRRRRHT 60
Db 1 MSDGTAEKSGGTG--FLSPAPAPKNGSSSDSSVGBKLGAADAVTGRTEYRRRRHT 58
QY 61 MDKDSRGAATTTTTHRRFRFRSVCDSNATALEPLPLSTLPSPSPAAPVQAGAPPEH 120
Db 59 MDKDSRGAATTTTTHRRFRFRSVCDSNATALEPLPLSTLPSPAAPVQAGAPPEH 118
QY 121 REETVATATSOVAQOPPAAPAGAEQAVAPASTVSSSTKRPVSGPSLVSKEPPP 180
Db 119 REETVATATSOVAQOPPAAPAGAEQAVAPASTVSSSTKRPVSGPSLVSKEPPP 178
QY 181 AR--SSGSGSAKEPQERSQOQDDIELETKAVGMSNDRFLKFDIEIRGSKFTYK 238
Db 179 SRSGSGSGSAKEPQERSQOQDDIELETKAVGMSNDRFLKFDIEIRGSKFTYK 236
QY 239 LDLETTVEAMCGLQORKLTTSERORFKEAEMLKGLQHNIVYRFDSSWSTYKGGKCI 298
Db 239 LDLETTVEAMCGLQORKLTTSERORFKEAEMLKGLQHNIVYRFDSSWSTYKGGKCI 298
QY 239 LDLETTVEAMCGLQORKLTTSERORFKEAEMLKGLQHNIVYRFDSSWSTYKGGKCI 298
Db 239 LDLETTVEAMCGLQORKLTTSERORFKEAEMLKGLQHNIVYRFDSSWSTYKGGKCI 298
QY 299 LVTELMTSGTLKTYLKRFRVMIKIVLRSCROQLKGLQFLHTRTPLIHDLKCDNIFIT 358
Db 299 LVTELMTSGTLKTYLKRFRVMIKIVLRSCROQLKGLQFLHTRTPLIHDLKCDNIFIT 358
QY 359 GPTGSVKIDGLGATLKRAKSPKSVTGTPEFNAPEMYEKKYDESVDVYAFGCMLEMAT 418
Db 359 GPTGSVKIDGLGATLKRAKSPKSVTGTPEFNAPEMYEKKYDESVDVYAFGCMLEMAT 418
QY 419 EYPSSECNAQIYRRVTSQVAPASFDKVAIPVKKIIEGCTRONDEKYSTIDLNHAF 478
Db 419 EYPSSECNAQIYRRVTSQVAPASFDKVAIPVKKIIEGCTRONDEKYSTIDLNHAF 478
QY 479 FQETGVAVELAEEDDEKIAIKLWLRIEDIKLKAGKYDNALIESFLERDVPEDVAQ 538
Db 479 FQETGVAVELAEEDDEKIAIKLWLRIEDIKLKAGKYDNALIESFLERDVPEDVAQ 538
QY 539 EHWESGYVCEGDHKTMAKIKRVSILKRRQROLVREOEKKQOESSSLKQOYE-QSS 597
Db 539 EHWESGYVCEGDHKTMAKIKRVSILKRRQROLVREOEKKQOESSSLKQOYE-QSS 597
QY 598 ASQTGIKOLPASASTGIPTASTTSASVSTQVEBEPEADQHOLOQVQPSISVLSGTYDS 657
Db 598 ASQTGIKOLPASASTGIPTASTTSASVSTQVEBEPEADQHOLOQVQPSISVLSGTYDS 657
QY 658 GQSSSVFTESRVSSQQTYSYSGOHEQANSTGTVPGHIPSTVQAGQSPGAVYPPSSVAQOQ 717
Db 658 GQSSSVFTESRVSSQQTYSYSGOHEQANSTGTVPGHIPSTVQAGQSPGAVYPPSSVAQOQ 717

QY 718 SOGQSSSSLTGVSSSQPIQHPQOQGGIOQTAPPOQTVQYLSQTSSEATTAQPVSPQ 777
 Db 719 NOGQP-SSSLAGVLSQVQHP-QOQGIPTVPPOQAVQYSLPQAASSSEG-TVQPVSQ- 774
 QY 778 QAPVULPVQYSAQGO-STQGVSGVAPAEVVAVQPCATOPTLLASSVDSASHVSGMSDG 836
 Db 775 -----PVYSAQTQSTQGVSGVAPAEVQPIPIQSQGTQVPLVSSVDSASHVSGMSDG 828
 QY 837 NENVPSSGRHGRRTTKRHYKSVNRSRHEKTSRKRLILVNSKGDVVECOLETHNR 896
 Db 829 NENAPSSSRHGRRTTKRHYKSVNRSRHEKTSRKRLILVNSKGDVVECOLETHNR 888
 QY 897 KWTTFEFDLDGDNPEELIATIMVNDLILAIERESFVDQVREIIEKADEMLSECVSEYEPG 956
 Db 889 KWTTFEFDLDGDNPEELIATIMVNDLILAIERESFVAQVREIIEKADEMLSECVSEYEPG 948
 QY 957 DQGLJESLQGDYGFSGSQKLEGEFYQPIPASSMPQOIGITSSLTQVHSAGRFIVAP 1016
 Db 949 DQGLJESLQGDYGFSGSQKLEGEFYQPIAVSSMPQOIGITSSLTQVHSAGRFIVAP 1008
 QY 1017 VPESRLRESKVFPSSEITDTVAASTAOSFGMNLSSHASLSLQAFSELRACMTTEGPNTA 1076
 Db 1009 VPESRLRESKVFPSSEITDTVAASTAOSFGMNLSSHASLSLQAFSELRACMTTEGPNTA 1068
 QY 1077 PPMFHTGTFEPVPPPLSSLAGVPTTAANT--APVATSSPNDISTSVIOSEVTPTE 1134
 Db 1069 PPMFHTGTFEPVPPPLSSLAGVPTTAANT--APVATSSPNDISTSVIOSEVTPTE 1125
 QY 1135 EGIAGVATSTGVVTSGLPIPVSESPVSSVSSITIPAVSISTPSLQVPTSTET 1194
 Db 1126 KGIQGVATSTGVVTSGLPIPVSESPVSSVSSITIPAVSISTPSLQVPTSTET 1183
 QY 1195 VWSSTALYPSVTVSATSASAGSGTATGPPVPAVVSQAAAGSTTGAITLTSVSTTSPS 1254
 Db 1184 IASSTGSPGTFSTTGTTVSSVAVANAPKPTVTLQOVVGNAGVAVIVSVSTTTPFA 1243
 QY 1255 TASQSLTQSSSTPPLATVTVVVAHSLDKTSHSSSTGALFSLAASSSSSPGAGVSY 1314
 Db 1244 MASQSPPLSSSTAPPLATVTVVVAHSLDKTSHSSSTGALFSLAASSSSSPGAGVSY 1303
 QY 1315 ISQPGAHPIVPIPSVIASTPILPQAAQPTSTPLIPVPSIPPLVQVPAVPAVQOULIHS 1374
 Db 1304 VSQGIYHPIVPIPSVIASTPILPQAAQPTSTPLIPVPSIPPLVQVPAVPAVQOULIHS 1363
 QY 1375 QPQPAALPNQPHTCPEVSDTOPKAPGIDDKTLEKRLASLBESHSAGQAHSVLET 1434
 Db 1364 QPQPAALPNQPHTCPEVSDTOPKAPGIDDKTLEKRLASLBESHSAGQAHSVLET 1423
 QY 1435 SLVIESSTYPTGIFTTAAVASKLSTSTSTCLPPTNLPLGTVALPVTPVTPFGQVSTP-- 1491
 Db 1424 PLVVE-IVTGPITPTTAAVASKLSTSTSTCLPPTNLPLGTVALPVTPVTPFGQVSTP 1482
 QY 1492 --VSTTSGVKGPGTAPSKPPLTAPVLPVGTTELPAITLPESEOLPPPPGSLTQSQPLE 1548
 Db 1483 ASAPASTITGAKPGTTPPKPLTITKVPVPGTELSAGVPCEQPLPPFGPSLITQOQPLE 1542
 QY 1549 DLDQALRTLSPEKITVTSVAVGVPVMAAFLAITTEAGTQPGQGVQYKEGFLAASSGAGV 1608
 Db 1543 DLDQALRTLSPEKITVTSVAVGVPVMAAFLAITTEAGTQPGQGVQYKEGFLAASSGAGV 1599
 QY 1609 FKMGHFOVSVADAQAGKXKNSSEDAKSVHESSTSESVSSSSPESTLVKPEPNCIT 1668
 Db 1600 VKMGHFOVSVADAQAGKXKNSSEDAKSVHESSTSESVSSSSPESTLVKPEPNCIT 1659
 QY 1669 FGISDVESAHKTTASAKSDTQGPLVGRFOVTTTANKVGRFSVSTKEDXITDTKXEG 1728
 Db 1660 SGISLDVDSHRTTTPAKSETQGPLVGRFOVTTTANKVGRFSVSTKEDXITDTKXEG 1719
 QY 1729 PVASPPMDLEQAVLPAVTPKKEKELSEPHLNKPSDDPEAAFLSRVDQSGSPHSPH 1788
 Db 1720 PVTS--PFYDSQEQTVLPAIPKKEKELSEPHLNKPSDDPEAAFLSRVDQSGSPHSPH 1778

QY 1789 QLSKSLPSQNLQSLSSNSFNSSVMSNDNESDIEDDKLETLRLRDLKELKEIOLQSRQ 1848
 Db 1779 HLCSSLEPIQLTQSLSSNSFNSSVMSNDNESDIEDDKLETLRLRDLKELKEIOLQSRQ 1838
 QY 1849 KEIETSLTKLGKVPVAVIIPPAFLSGRRRPTSKGSKSSRSSSLGKNSPOLSGNLSG 1908
 Db 1839 KEIETSLTKLGKVPVAVIIPPAFLSGRRRPTSKGSKSSRSSSLGKNSPOLSGNLSG 1898
 QY 1909 QSAASVLPQOQLHPGPNIPESQNLQPLXPSPSSDNLVSAFTSDGAISVPSLSAPQ 1968
 Db 1899 QSGTSLVLANQOQLHPGPNIPESQNLQPLXPSPSSDNLVSAFTSDGAISVPSLSAPQ 1958
 QY 1969 GTSTNTVATVNSQAQAQAPPAATSSRRKGTETDILHKLVDWARDAMNLSGRGSKGM 2028
 Db 1959 GTSTNTVATVNSQAQAQAPPAATSSRRKGTETDILHKLVDWARDAMNLSGRGSKGM 2018
 QY 2029 NTEGPMARKFAPQGLCTSMNSNGGSPISAASTSLGHTTKSMCPQOYGFPAAPG 2088
 Db 2019 NTEGPMARKFAPQGLCTSMNSNGGSPISAASTSLGHTTKSMCPQOYGFPAAPG 2078
 QY 2089 AQMSGTGAPAPQPLQGFQVGTASLQNFNISNLQKSIINPSPSNLRTT 2136
 Db 2079 TQMSGTGAPAPQPLQGFQVGTASLQNFNISNLQKSIINPSPSNLRTT 2126
 RESULT 4
 Q9PIS9 PRELIMINARY; PRT; 670 AA.
 AC Q9PIS9;
 DT 01-OCT-2000 (TEMBLrel. 15, Created)
 DT 01-OCT-2000 (TEMBLrel. 15, Last sequence update)
 DT 01-OCT-2003 (TEMBLrel. 25, Last annotation update)
 DE Kinase deficient protein KDP (Fragment).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Breast carcinoma;
 RX MEDLINE=20127920; PubMed=10660600;
 RA Moore T.M., Gary R., Johnson C., Coptrout M.J., Ridley A.J.,
 RA Morris J.D.H.;
 RT "PSK, a novel STE20-like kinase derived from prostatic carcinoma that
 RT activates the JNK MAPK pathway and regulates actin cytoskeletal
 RT organization.";
 RL J. Biol. Chem. 275:4311-4322(2000).
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 DR EMBL; AF061944; AAF31483.1; -;
 DR GO; GO:0005524; F:ATP binding; NAS.
 DR GO; GO:0004674; F:protein serine/threonine kinase activity; NAS.
 DR GO; GO:0006468; F:protein amino acid phosphorylation; NAS.
 DR InterPro; IPR006719; Prot kinase.
 DR InterPro; IPR008271; Ser_Thr_kinase.
 DR Pfam; PF00069; Pkinase; 1.
 DR ProDom; PD000001; Prot kinase; 1.
 DR PROSITE; PS00101; PROTEIN KINASE DOM; 1.
 DR PROSITE; PS00108; PROTEIN KINASE ST; 1.
 KM ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
 FT NON TER 670 670
 SQ SEQUENCE 670 AA; 73695 MW; C24A9A8D7FC32AA CRC64;
 Query Match 31.5%; Score 3408; DB 4; Length 670;
 Best Local Similarity 100.0%; Pred. No. 6; 1e-159;
 Matches 668; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MSGAAEKQSTPSGLSPSPAPAPKXGSSDSSVGEGLGAAADAVYGRTEYRRRHT 60
 Db 1 MSGAAEKQSTPSGLSPSPAPAPKXGSSDSSVGEGLGAAADAVYGRTEYRRRHT 60
 QY 61 MDKDSRGAATTTTTHFRFRRSYICDSNATALELPGLELPPSPIPAAVQSAAPPBP 120
 Db 61 MDKDSRGAATTTTTHFRFRRSYICDSNATALELPGLELPPSPIPAAVQSAAPPBP 120

DR GO; GO:0004713; F:protein-tyrosine kinase activity; IPA.
 DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
 DR InterPro; IPR000719; Prot kinase.
 DR InterPro; IPR002290; Ser_thr_kinase.
 DR InterPro; IPR008271; Ser_thr_kinase.
 DR InterPro; IPR001245; Tyr_kinase.
 DR ProDom; PD000001; Prot_kinase; 1.
 DR SMART; SM00220; S_TKc; 1.
 DR SMART; SM00219; TyKc; 1.
 DR PROSITE; PS00011; PROTEIN KINASE DOM; 1.
 DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
 DR Hypothetical protein.
 FT NON TER 1
 SQ SEQUENCE 2219 AA; 234470 MW; B7AD70196CZA92C0 CRC64;
 Query Match 22.2%; Score 2615; DB 4; Length 2219;
 Best Local Similarity 31.9%; Pred. No.2.4e-119;
 Matches 811; Conservative 234; Mismatches 645; Indels 856; Gaps 86;
 QY 60 TMDKSRGA-----AATTTTTHFRFRSVYICDSNATALEPLSLPQSPSTPAA 110
 DB 8 TLMPRGAGAGPAGMAEPRAKAAAPGQRFILRSV---ESQDEHPGLF-----AAE 56
 QY 111 VPQAPPEPRHEE-----VTATTSQVAAQPPAA-----AAPGEQAVAGPASTV 156
 DB 57 APGPQPPQPLQRRVLLICKTRRLTAERARGRPAAPAPALVAOPGAPADAGPEP--- 113
 QY 157 PSSTSKRPVSPSLVSGSKPEPP-----ARSSGSGSASAK-EP-----QERSQ 199
 DB 114 ---VGTEPBPDPFLAAVETNAPADGGRPEAAATVRKEDEGAAPPEPRTTRDEE 170
 QY 200 QQDDIEELTKAVAGMSNDGRFLKEDIEIGSGSFYTKVKGDLTETTVAVMCELODRLT 259
 DB 171 EEDDEDL--KAVATSLDGRFLKEDIEIGSGSFYTKVKGDLTETTVAVMCELODRLT 228
 QY 260 SEORFEAEEMKLGLOHNPVIRPYDSESTVGKCCIVLVTIELMSTGLTKYLKRPKM 319
 DB 229 LERQRFEEEMKLGLOHNPVIRPYDSESTVGKCCIVLVTIELMSTGLTKYLKRPKM 288
 QY 320 KIKVLSMWCQILKGLQFLTRPTPIIHRDKCNFIITGPTGSVKIGDLGLTKKASF 379
 DB 289 KRPVLSMWCQILKGLFLTRPTPIIHRDKCNFIITGPTGSVKIGDLGLTKKASF 348
 QY 380 AKSVIGTPEFMADMEYKYDESVDYVAFGCMLEMAISEPYSECMAQIYRYATSGV 439
 DB 349 AKSVIGTPEFMADMEYDESDYVAFGCMLEMAISEPYSECMAQIYRYATSGV 408
 QY 440 KPAKSPDKVAIPYEKIEIGGIRQNKDERYSIKOLLNAFFQSEETGVAVELAEEDGSK- 497
 DB 409 KPAKSPDKVAIPYEKIEIGGIRQNKDERYSIKOLLNAFFQSEETGVAVELAEEDGSK 468
 QY 498 IAIKMLRIEDIKLKGKYKDNFAIFSPFLERDVEDVAQOEAVESGVEGSHKTKMAA 557
 DB 469 IAIKMLRIEDIKLKGKYKDNFAIFSPFLERDVEDVAQOEAVESGVEGSHKTKMAA 526
 QY 558 IKDVSLLIKRKQRQRLVREDEKKQOESSLKQOYEOSSAQGTGIKQLPSASTGIPAS 617
 DB 527 IKDVSLLIKRKQRQRLVREDEKKQOESSLKQOYEOSSAQGTGIKQLPSASTGIPAS 580
 QY 618 TTSASVSTQV-----EPREPADHQOQLOYOQPSI-----SVLSGTVDSGGSSVTE 666
 DB 567 QVQVYTHAAGGPGPEPEPEPADQ-----LAPTLPTSATSLASTSTDSGGSTVSD 622
 QY 667 SRVSSQQTIVSYGQHEQAHTSG-----TVPGHLPSTVVOAQSPHG----- 706
 DB 623 SQ--SSQGSQVMSLGLAADAASPAQCVCSPVSSQPVLPQSLPS--LGAVQDPALPGLPVSS 680
 QY 707 -----VYPPSSVAQSGSGSPS----- 723
 DB 681 VPAPACPSPLQHPDPDPAWSFAPVLPBPSTPMTGQGPAPGQGPPLAQPTPLPQVLA 740
 QY 724 -----SSSLGVSSQPTQHPQDQ--QGIQQ----- 747

DB 741 PQPVVLPQVPPPHLPYLAASQVGAAPQLKPLQMPQAPLOPLAQVPPQMPPIPVVPEIT 800
 QY 748 -----TAPPOQTVOY----- 757
 DB 801 PLAGIDGLPALPDLPLATVPPVPPPPQYFSPAVILPSLAALPSPASPLPQAVLPHRP 860
 QY 758 ----- 757
 DB 861 GAPLAMPCKRTIVNAPATIPLLAVALPQVALISHSAAVALPQGVVPPAEPQMAPTDVP 920
 QY 758 -SLQSTSSSEATTAOPVSOQ----- 778
 DB 921 PSRHHTYQNMKATPQPALPQPPQPTPLPQPVLPQPTLPPEQVLPPEQTPPQVLPQPM 980
 QY 779 ----- 778
 DB 981 LPQPVLPQVPPHLPVREPLQPHLPQGAAPATPGSQIILGHAPAVANDVAQVPTVVP 1040
 QY 779 -----APVLPQ-----VSAGKQ--TQGVSAVAP----- 802
 DB 1041 PAAVSPPLPBEVLLPAPBELLPQPPSLATVASAVQSVPTQTATILPAPNPLPGGPGA 1100
 QY 803 -----EPVAVAQPOATOPTTLASVDS--ASHVVASG--MSDQENVPSSSGRHEG 849
 DB 1101 SPCTVQLTEPVOEBAQSDKPRGLPQSCSYGSDVTSKELSDSCGA--FGGRLG 1159
 QY 850 FTKRHYRKVSRSRHEKTSRPLRLILVNSKGRVVECOLETHNRKMTFKFDLDGN 909
 DB 1160 RAARKHRRSTFARSROERASRPRLTILVNCMDKWKVEQLEHNNKMTFKFDLDGDA 1219
 QY 910 PEEIATIMVNDPILAIERSFPVDQVREIEKADENISEVSVPEEDQGLE---SLQG 965
 DB 1220 PDLATYVMEHDILQERETFLQMKDVNDKADMLSEDTLDKRGSDPGTSPHLSGTC 1279
 QY 966 KDDYGFSGQKLEGEFKOPIPASMPQOIGIPSSILQVHSAGRRTVSPVESRLRES 1205
 DB 1280 LG---TGESSRSQANAPV---YQ-----NVHTGKRWITICPVNHPPEA 1321
 QY 1026 KVPSEITDTVAASTQSPGMLSHSASSISLQAFSELRRAQTEGPNTAPNFSTHP 1085
 DB 1322 ---PESSPPLPSSLPEASQDSAPYKQDLSKQPSFLASQDL-----LSQAGP 1368
 QY 1086 TFPVVEPFLISLAGVYTTAATAFPV-ATSSPPNDISTSVYQSEVTVPEEGIAGVATST 1144
 DB 1369 SNP-----PGAP-----PAPLAPSSP-----VVALPQDG--AAATST 1400
 QY 1145 GVVTSGGLPIPPVSESPVLSVVSITIPAVVAGISTTSPSLQVPTSTSEIVVSTALYPS 1204
 DB 1401 ---MEP----- 1404
 QY 1205 VIVSATSSAGGSTATPGKPPAVVSQAAGSTTVGATLTSVSTTSFSTASQLSIOIS 1264
 DB 1405 --ASGTASQAGG--PG-----TPQGLTSELTQPLAETHEAPLAVQPIVVGLA 1449
 QY 1265 SSTSTPFLAETVWVSAHSLDKTSHSSTGLABSLAPSSSSSGAGV--SYISQPGGLAP 1323
 DB 1450 PCTPAPEASTDASAP-----RBPDPAPPEPSHSGTPPQALQPPAPLP 1496
 QY 1324 LVIPSVIATSTPLPQAAGPTSTPLPQVPSIPPLVQPVANVPVQQTLLHSQPPALPN 1383
 DB 1497 AAVGAVSLATSQLP-----SPPLGPTV-----PQPSAL--- 1526
 QY 1384 QPHTHCPEVDSTQPKAGID--IKTLEKLR--LFSHSSGGAHASVLSLTVIRS 1440
 DB 1527 -----ESDGGGPPPVGVFVDSITIKSLDEKIRTLLYQHVPVTSASAGTPEV----- 1573
 QY 1441 TVTPGIPTTAVAPSKLILSTSTCLPPTMLPIGTVALPTTP--VTPPGQVSTVSTTSGV 1499
 DB 1574 ---GDRDFTLEPLKDGQPRSEVC-----GADLALPVPPEAVSAGRQLP----- 1614
 QY 1500 KPGTAPSKPPLTKAALVY--GTLEPAGTLPSBQLPPFPSPSTQSQPFLDY--DAQLR 1556
 DB 1615 -----QPIVEKSLAIPTRGAVMEQGT-----SSSHTAASSPRSMGLGYDRGQ 1657


```

QY 910 PEEATATWNNDFILAIERESFVDQVREIIEKADMLSEEDVSVEPEGQJESLQKXDY 969
Db 782 PEDIADYVVENENFVLESEKEFEVBEIARIVGQAQIL--HHAFERARATGDSITVDSNS 839
QY 970 GFSGSQKLEGEFKQCIIPASSHPQOIGIPTSSLTQVHSHAGRRFIVSPVESELRRESKVP 1029
Db 840 SQTSS-----BOVQINSTSTQTSNESAPOS--SPVGRWRF-- 874
QY 1030 SEITDTVAASTAQSPGMNLSHSSASLSIQAFSELRARRQMTGEPATAPNNSHGTPTFPV 1089
Db 875 --INQITRNRETQS-----PPSLQH-- 892
QY 1090 VPPPLSIAVPTTAATAAPVAPATSSPPN-DISTSVIOSEVTPTEEGIAGVATSTGVYT 1148
Db 893 -----SMAVAP-----GRHPLPSPKNTSNKEISNDTLTINNPOHRLFTSKSEHKDQV 942
QY 1149 SGGPIRPVSESPVLSVSSITIPAVVISITSPSLQVPTSTSEIV--VSTALYPSV 1205
Db 943 DG-----KISECAVETKQPAIYQVEDNRQJMAPVTNSSYSTT 982
QY 1206 TVSATASAGGSTATPGPKPRAVVSQQAAGSTVGAITTSVSTTSPSTASQJSLSS 1265
Db 983 SVRAVPAECBELTQASIFIPVPCHQ-----TASQADALMSHPESTQ----- 1026
QY 1266 STSTPTLAETVVAHSLDKTSHSSTTGLAFSLAPSSSSPGAGVSYSIQPGALHPLV 1325
Db 1027 -----TSGNSLTITIAFDQKQPTLSVQPGAMDHEFISQES----- 1060
QY 1326 IPSVIATSTPLPQAAGTSTPPLPQVPSIPPLVVPVAVVAVOQTLHSGOPALPLNOP 1385
Db 1061 -ETVNTBASSPKVITPQTQGLE-----PTTLOP----- 1089
QY 1386 HTNCPVEDSDTOPKAPGIDD-IKTLFEKLRS-LFSEHSSSG--AQHASVLETSIVIE 1439
Db 1090 -TTLVBSDBRPRPLEFADNRKITLDEKLNLYQEHISISIPESKQJQSDSPSSS 1148
QY 1440 STYTPGIPITAV-----ABSKLLTSTTCLPPTNPL----- 1472
Db 1149 AEDTLSCPTVEVIAISHGQIKDSPVQSPNFOQTSKILSNVAAS--QPANISVFKRDLNV 1206
QY 1473 -----GYALPVTVPVTPGQVST--PVSTTSSGVKPG--TAPG-KPEPLTK 1512
Db 1207 ITVSPESELCHENMSDASLPGDPEAVPAVSSGGAHILQTGGGFGISFTCPSLKPNISK 1266
QY 1513 APVLPVGTBLPAGTLPESEQLPPPPSPSLTQSQPLEDLDAQLRRTS-----PEX 1563
Db 1267 -----KSWIRK--LKSMAVRLQOSTSFFKRSVROVETE 1298
QY 1564 TVTSAVGPVSMAPTAITEAGTOPKGVQKGPVLAATSSGAGVFMGRFOVSVADGA 1623
Db 1299 EMMSATAP--DPILTRESTDTRALNRCK-----AMGSPQORGFQYITTPQ-- 1344
QY 1624 OKRGGKKSSEDASVHSESSSESYSVS--SSPSTLVKPEPNITTPGISSDVPEBAHKT 1682
Db 1345 QQSAAKTSRGIEHTSVFSETNHSSEAFIKTAASQVEIEP----- 1385
QY 1683 TASEAKSDTGQPTKVGKFOVTTANKVGRFSVSKTEDKITDTCKEG--PVASPPMDLE 1739
Db 1386 -----ATQNKITSFSYKLO-ALQETKEKGVKPKQGDNPLSFS 1423
QY 1740 QAV--LPVAPIPKKEPE-----LSBPSHLNG--PSSDPE-AAFLSRDV 1777
Db 1424 AACETDVSVTPEKEFEFTSATGSSWQSGSELLKREILITAGKQSSDSFSSASIA-- 1480
QY 1778 DDGSG-----SPHS--PHQSSKSLSSQNSQSLNSF--NSGYMSDNESDIEDD 1825
Db 1481 --GSGKSAVAKTGPESNOCLPHH-----EQQVAAQTQSGSLFYSPPSPMSDDSEIIEED 1532
QY 1826 LKLELRRLDKLKEIKIODLSROKHEIBSLYTKLGKVPVPA-----VIIPAPALSGRRR 1879
Db 1533 LKVELQRLREKTIQEVNLTQOONKELQYERLMSIKRSKQSTQSTIEPIPLPASP-----R 1587

```

```

QY 1880 RPTKSKSKSSRSSSLGNKSPOLSGNLSGQSAASVTHPQUTLHPGNIPEGQNLQPL 1939
Db 1588 RRSFESKSLRSR-----POSILTHVDNGIVAGKSLINEL 1622
QY 1940 KPSPSDNLYSAFTSDGAI SVPSLSAPGQSTSTNTVATVATNSQAQAQAPAMISSRGT 1999
Db 1623 E-----NPLCVESNMAASCQSP--ASKRGM 1645
QY 2000 FTDDLHLKLVNDWARDAM-----NLGRRGSKGHMNYEGPMARKFAPQOL-----CI 2047
Db 1646 FTDDLHLKLVNDWTKRKAAGVNSLTKPSLNLQKQSKHLEFEMNNKXSE-NTPEMTGTSWTI 1704
QY 2048 SMTNLSGSAPIASAASATSLGHFTKSMCPPOQYGP-----ATPEGAQMSGTGPA 2098
Db 1705 SLSLQIRGAFTSLPQGLSLPSFPG--PLSSYGMFHVQYNAVAGAGYPQWVGISGTT 1761
QY 2099 POPU-----GQPPVGTASLQNFNTSNLQKSIJNPG 2130
Db 1762 QQSVVIPAQSGGFPQ--GMMQAFPTTSVQNPATITPG 1798

```

RESULT 8

```

QY 09BYP7 PRELIMINARY; PRT; 1743 AA.
AC 09BYP7
ID 01-JUN-2001 (Tremblrel. 17, Created)
DT 01-JUN-2002 (Tremblrel. 21, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE Protein kinase MNK3.
GN PRKNK3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=21455683; PubMed=11571656;
RA Verissimo F., Jordan P.;
RT "MNK kinases, a novel protein kinase subfamily in multi-cellular
RT organisms."
RL Oncogene 20:5562-5567 (2001).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Jordan P.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
CC -! SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; AJ409088; CAC32455.2; -.
DR GO; GO:0005524; P:ATP binding; IEA.
DR GO; GO:0004674; P:protein serine/threonine kinase activity; IEA.
DR GO; GO:0016740; P:transferase activity; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR000719; Prot kinase.
DR InterPro; IPR008271; Ser Thr_pkin_AS.
DR Pfam; PF00069; pkinase; I.
DR ProDom; PD000001; Prot_kinase; 1.
DR PROSITE; PS0011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 1743 AA; 191788 MW; 7B06C8C7B8E4D57 CRC64;

```

Query Match

Best Local Similarity 33.6%; Score 2439; DB 4; Length 1743;

Matches 746; Conservative 246; Mismatches 617; Indels 610; Gaps 80;

```

QY 27 NGSSSDSVGEGLGAAADAVGTGTEERYRRRHNTMDKSRGAATTTTHERFFRRSYIC 86
Db 18 DGISEFNRVPO---VAATTTVEARLKE-----KSTFSASGETVERKRFKSV-- 63
QY 87 DSNATALELPLGLPLSLPQPSIPAAVPOSAPPEPHREFTVATATQVAAQPPAAAPGQ 146
Db 64 -----EMTEDDKVAESSP-----KDEKIKAAAMNIPRVDKLPNSVLRGQ 102

```


147 AVAGPASTVPSSTKDRPVSOPLVSGXKEPPPARSGSGSGAKEPOEERSQOQDDIEE 206
 103 EKVYECOSSTISISD-----CFKKNKEME-----BE 132
 207 LETKAVGMSNDGRFLKEDIEIGRGSFKYKGLDTEITVEVAMCELQDRKLTSEFORIK 266
 133 AEMKAVATSPSGRFLFEDIELGARGFKTYKGLDTEITVEVAMCELQDRKLTAEQORER 192
 267 EEMKMLKLOHPNIVAFYDSMESTVNGKCTIVLTELMTSGTLKYLKPKMKIVLVS 326
 193 EEMKMLKLOHPNIVAFYDSMESTVNGKCTIVLTELMTSGTLKYLKPKMKIVLVS 252
 327 MCGOIKGLQLOFHTPTPIIHRDLKCDNFITPTSGSVKIGDLATLKRASFAKSVIGT 386
 253 MCGOIKGLQLOFHTPTPIIHRDLKCDNFITPTSGSVKIGDLATLKRASFAKSVIGT 312
 387 PERMAPEMEYKEDSVYVAFGCMLEMTSEYVSECONAAQIYRVTSGVKPASFDK 446
 313 PERMAPEMEYKEDSVYVAFGCMLEMTSEYVSECONAAQIYRVTSGIKPASFMK 372
 447 VAIPVEKEITEGTRQKDRYSIKDILNNAFQOEIGVAVLEEDD--GEKIAIKML 504
 373 VTDPEVEKEITEGTRQKDRYSIKDILNNAFQOEIGVAVLEEDD--GEKIAIKML 431
 505 RIRIDIKLKGKYNDNEAIEFSFDERVDPEDVAQEMVSGVCEGDHKTAKAKIKRVS 564
 432 -VEDPKLKGKNDNEAIEFSFNLDTDPREVAIEMKSFHESDSKAAPKIRDRVTP 490
 565 IKRREOROLVREOEKKEESSLKQOVSOSSAQIGIKOLPSASTGIPASTTSASVS 624
 491 IKRRE-----KKPAGCLERRDSQCKSMGNVFPQPQ--NTTLPLA-----PAQ 533
 625 TOVEPEPEADQ--QOLQYQOBS--ISVLSGTVDSGQSSVFTTSRVSQOQTVYGS 679
 534 TGACEETEVDOHVRQOQLQKPKQOCHSVTGDMLSEAGASVHSDT--SSQSVAYSS 591
 680 QHEBASHTGVPHGIPSTV--QAO--SOPHGVYPPSSVAGQSQOQPSSSLTGVSSQPIQ 737
 592 NO-----TWGSOVMNIPQAEVNVFPQIT-----SSQQLVG 622
 738 HPQOQOQIOQTAPRQOQTVQYSLQSTSSSEATTAQVSPQAPQVLPQVSAKQOSTQVS 797
 623 HYQVSGIQKHSKLTQOQILPLVQSQ-----STVLPLV--HVLGPVTVQPVPLTVQVP 676
 798 QVAP--ABPVAVAQOQATOPTTLASSVDSASHDVASGMSDGNEN-----PSSGRHGR 851
 677 QIKVVSQPVGAEOQOAAALIKPDL--VRSINQDVAT--TKENVSPPDNPSGNGKODRIK 729
 852 TKRHYKRSVRSRHEKTSRPKLRIINVSNGKDRVVEGQLETHNRKMYTFFEDLDGNPE 911
 730 QRR-----ASCPRPEKGTFOQLTVLOVSTISGDMWVEQLETHNRKMYTFFEDVDGAPR 783
 912 EIAITMVNNDIILAIERESFVDOVREIEKADMLSEDSVVEPBGDQLESLOKDDYGF 971
 784 DIADYMWEDNVLSEKKEKFEELRAIVGQOELI--HVHFATERATGVDSITVDSNSQ 841
 972 SGOQKLEGEFKOPIPASSMPOQIGIPSSLTQOVHASARRFIVPVPYSRLRESKVPSE 1031
 842 TGSS-----EOVQINSTSTQTSNESAPQS--SPVGRMRC-----874
 1032 ITDVAASSTASPGMNLHSHASLSLQOAFSELRAQCTEGPNTAPREFSTGTFPVVP 1091
 875 INQITRNRETS-----PESLOH-----892
 1092 PFLSIAGVPTTAATAVAPATSSPPN--DISTVQISEVTVPTBGLIAGVATSTGVNSG 1150
 893 ---SMSAVP---GRHPLSPKNTSNKEISMDTLTIENNDCHRALFTSSEHDDVDG 944
 1151 GLPIPVESPVLSVSVSITTPAVVSIISTSPSIQVPTSISEIY---VSTALYPSVTY 1207
 945 -----KISBCASVETKOPAILVOYEDNRQIMAPVNTSSSYSTVS 984

1208 SATSASAGSTXITPCKPRAVAVSQOAGSTTVGATLTSVSTTSPSTASQLSLOLSST 1267
 985 RAVPAECBGLTKQASIFIPVPCHQ-----TASQADALMSHFGESTQ-----1026
 1268 STPTLAETVWASAHSDKTSSTTGLAFSLAPSSSSPPGAVSSYISQPGIADPLVLP 1327
 1027 -----TSGNSLTLTADQKRPQTLISVOQPMADAFIQOB-----E 1061
 1328 SVIASTPILPOAGPTSTPLPOVPSIPPLVQPVANVAVQOTLHISQOPALLPNOPT 1387
 1062 TVNTEASSPKTIVIPQTPGLE-----PTTLQP-----T 1090
 1388 HCEPVDSDTOPRAGIDD--IKTLEKLS--LSEHSSSGAQHASVSLSTVIESTVTPG 1445
 1091 TVLESDEGRPPKLEFADNRIKTLEKRLMLQOEHIS-----STPESQ-----1135
 1446 IPTTAVAPKLLTSTSTPLPPTNLPLGVALPVTPVTV--PGQVSTPST-----TTS 1497
 1136 -----KOTISDPSFSSSADTLSCVTVEVIAISHCGIKDSFYQSPFPQOTGS 1183
 1498 GVKPGTAPSKPP-----LTKAPVLPVGTLEPAOTLPSE--QLP--PEPGLTOS 1543
 1184 KLSNVAASQPANISVFKRDLNVITSV-----SELCEHMSDASLPGDPEAYPAVSS 1238
 1544 QOPLDLDQRLRTTSPKXITVTSANGVPSMAAPTAITBAGTOPQKQVSGVKEGPVLATS 1603
 1239 -----GGNHLQTVETHEKMSATAP--DPIPLRESTDALNCK-----1279
 1604 SGAGVFKMGRFOYVVAADAGQKXKSEDAKVHSESTSESSVLS--SSPSTLVKPE 1662
 1280 AMSGSQKGRFQVITTPQ--QCSAKMTSFGEIHTSVFSEIHNHSEEFITAKSOLVEI 1337
 1663 PNGITIPGISSDVPESAHKTJASBAKSDTQPTKVGAFQVTTANKVGRFSKTEDKIT 1722
 1338 P-----ATQNPKTSFSEKLO--ALQ 1356
 1723 DTKKSG--DVASPPMDLEQAV--LDAVTPKKEPE-----LSESHL 1761
 1357 ETCKENKGVKQGDNFIASFACETIVSVTPKEFEITSATSSMOSGELLKKEEIL 1416
 1762 NG--PSSDE--AAFLSRVDVDSG-----SPHS-----PHQLSKSLPSONLSQSLNS 1807
 1417 TAGKOPSSDEFSASLA--GSGKSVAKTGPESNOCLBHH-----EODAYAQTOSSL 1465
 1808 F--NSSYMSDNDSEIDEDLKLRLRDKHKEIDLOLSRQKHEIESLYTKLGVPVA 1865
 1466 FYSPSSPMSDDSEIEDLKLRLRDKHKEIDLOLSRQKHEIESLYTKLGVPVA 1525
 1866 ----VITPPAPLGRRRRPTKSKGSKSRSSSLGKSPOLSGNLSGQSAASVLAHQ 1919
 1526 KTQSTEIPLPPASP--RRPSFKSKLSR-----POS 1555
 1920 TLMHPGNIPESGQOQLQPLKPSPPSDNLVSATFSDAISVPSLASGQTSSTNTVGAT 1979
 1556 LTH-----VDN-----GIVADPLQVE 1572
 1980 VNSQAAQOPAMTSSKGTJTDILKLVDNMWARDAM-----NLSGRGSKGHMNYEG 2032
 1573 SNASCOOSP--ASKKGMFTDHLKLVDMWTEAVGNLSIKPSLNOLOKOSQKLETEN 1628
 2033 PGMAKTSAPGOL-----CISMTSNLGSAPISAAKATSLGHFTKSKCPPOQYGF--2083
 1629 WNKVSE--NTBSTWQVSTWISLSLQIKRQAVPTSPGUISLPSFG--PLSSYMPHYCQ 1684
 2084 ----ATPFGQWNGSTGPPAPOL-----GOFQPVGTASLQNFNISLQKSIINPPG 2130
 1685 YNAVAGAGYVQWNGISCTTOQSVITPAQSGGPPQF--GMMQAPFTSSVGNPATIIPG 1741
 996392 PRELIMINARY; PRT; 1243 AA.
 996392; 996DT8;

DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Putative protein kinase MNK4 (Serine/threonine protein kinase)
 DE (EC 2.7.1.37).
 GN PRKMNK4.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 OC NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Kidney;
 RA MEDLINE=21390047; PubMed=11496583;
 RX Wilson F.H., Disse-Nicodeme S., Choate K.A., Ishikawa K.,
 RA Nelson-Williams C., Desideri I., Gunel M., Milford D.V., Lipkin G.W.,
 RA Achard J.-M., Feely M.P., Dussol B., Berland Y., Thwin R.J., Mayan H.,
 RA Simon D.B., Farfel Z., Jeunemaitre X., Lifton R.P.;
 RT "Human Hypertension Caused by Mutations in MNK Kinases";
 RL Science 293:1107-1112(2001).
 RN [2]
 RP SEQUENCE OF 13-1243 FROM N.A.
 RC TISSUE=Kidney;
 RA Chistiakov D.A.;
 RL Thesis (2001), Department of INSERM U36, College de France, Paris,
 France.
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 DR EMBL; AF390018; AAK91995.1; -
 DR EMBL; AJ316534; CAC48387.1; -
 DR Genew; HGNC:14544; PRKMNK4.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0004681; F:GMP-dependent protein kinase activity; IEA.
 DR GO; GO:0004682; F:protein kinase CK2 activity; IEA.
 DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
 DR GO; GO:0016740; F:transferase activity; IEA.
 DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
 DR InterPro; IPR000715; Prot_kinase.
 DR InterPro; IPR008271; Ser_thr_kinase.
 DR Pfam; PF00069; pkinase; 1.
 DR ProDom; PD000001; Prot_kinase; 1.
 DR PROSITE; PSS0011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE; PSS0108; PROTEIN_KINASE_ST; 1.
 KM Hypothetical protein; ATP-binding; kinase;
 KW Serine/threonine-protein kinase; transferase.
 SQ SSQENCE 1243 AA; 134739 MW; BAC35F0938BA391 CRC64;
 QY Query Match 19.7%; Score 2126; DB 4; Length 1243;
 Best Local Similarity 31.9%; Pred. No. 9.8e-96;
 Matches 612; Conservative 191; Mismatches 393; Indels 724; Gaps 58;
 QY 10 SSTPGSLFSPAPAPKNGSSSDSVGEKLGAAADAVTGRTEYRRRRHMDKDSRGA 69
 DB 14 SQTEADLALRPPPLGTAGQ-----PRAGPP-----RRARFSK----- 49
 QY 70 ATTTEHFFRRSVYCDNATLALPLSLPQSPAPAVQAPPEPHEETATATA 129
 DB 50 AERPPSSRLSRSSVDLGLSSWSLPASP--APDPDP--PDSAGPBAR----- 96
 QY 130 TSOVAQPPAAPAGGAVAGPASTVPSSTSDRVSQPSLVGSKEPPPARSGSGGS 189
 DB 97 -----SPPSKSPPEPGTWTGAP--VKAEDSARPELPDSAVGPSREP----- 139
 QY 190 AKEPQE--ERSQODDIELELTKAVGMSNDGRFLFDIEIGSFYVYKGLDTEET 246
 DB 140 LRVPEAVALLERREGEKEDMETQAVATSPDGRVLFDEIEIGGSKTYRGLDITTYE 199
 QY 247 VAMCEIODKLTKESEQRKEKEEMKGIQHENTVAFYDSWSTVYKGCIVLVETLMT 306
 DB 200 VAMCEIQTTRKLSRAERQRESEVEMLKGIQHENTVAFYDSWSTVYKGCIVLVETLMT 259
 QY 307 GLTKTYLKEFKYWKIVLWSMCGQILKGIQFLTRPPPIIHDDIKDNFIITGPISVY 366
 DB 260 GLTKTYLKEFKYWKIVLWSMCGQILKGIQFLTRPPPIIHDDIKDNFIITGPISVY 319

QY 367 GDLGIATLKRAKSPAKSVIGTEPFNAPEMYEKEVDESVDVAFGMCLMAMTSEYPSBO 426
 DB 320 GDLGIATLKRAKSPAKSVIGTEPFNAPEMYEKEVDESVDVAFGMCLMAMTSEYPSBO 379
 QY 427 NAAQIYRVTSVGPVSPASFDKVAIPVKEITEGCIRONDERYSIKDLINAPFOSETGYR 486
 DB 380 NAAQIYRVTSVGPVSPASFDKVAIPVKEITEGCIRONDERYSIKDLINAPFOSETGYR 439
 QY 487 VELAEEDGKIAIKLWLRIDIKKLGKYKONEAIEFSFLEDPVDPDAQEWESGVY 546
 DB 440 VELAEEDGKIPKILWLRMEDAR--GGRPRDNGAIEPLFQGRNAAEVQAQEWALGLV 498
 QY 547 CEGHKIMAKIKRVSLIKRERQRLVREOEKKEQSESLQOQVSSAQSQTIRQL 606
 DB 499 CEADYQVAVRVRERVAIQKREKLRARE-----LEAL 533
 QY 607 PSASTGIPTASTTSASVSTGY---EPEPEADQHQQLQYQPSI--SVLSDGTVDGSGSS 662
 DB 534 P-PEPGPPRATVPMAPGPPSVFPEPEPEADQHQPLFLHASYSTTSDCTGVLSS 592
 QY 663 VFTESRVSSQQTWSYSGQHEQAHSTGTVPGHIPSTVQAQSPHGVPPSSVAQSQSQGP 722
 DB 593 GFLDASDPALQP-----PGGVPSL--ASHLCLESA----- 622
 QY 723 SSSSLTVSSSQPLQHEQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQO 782
 DB 623 -----FALS----- 626
 QY 783 LPQVSAQSGQSTQGVQVAPAEVVAQAQPTTLLASSVDSAHSDVAGSGDGNVPS 842
 DB 627 IPRSGPSSDPSG-----DSVADSAAGSLDVBGM-G 658
 QY 843 SSGRHSGRTTKRYKRSVRSRSHKTSRKLILNVSNKGRDVEVQCLTEHNRKMTFFK 902
 DB 659 QMRPPGKRNLR-----RPRSR-----LRVTSDDQNDRAVEQLQTHNSKMTFR 704
 QY 903 FDLDDGNPEIATIMANDITLIERESFVDQREITEKADENLSEPVSEPRGQGLS 962
 DB 705 FDLDDGSPFEEIAAMVYNEFLPSEBDGFLRIREITLQREVELTKRDTGMEAAEDTLSP 764
 QY 963 LQCKDYGFGSGQKLGSEFKQPIPASSMPQOIGIPITSLQVNVHSAKRRTVSPVPSRL 1022
 DB 765 QE-----EPAPLALVPLDPSNBEIQ-----SSTS 791
 QY 1023 RESKVPSEITDTVAASTAQSPGMNLHSASSLQOAFSELRAQWTEGPTAPVPSH 1082
 DB 792 LEHRSWTAFTSTSSSPOTPLSPG--NFSPEPTIPISGPFI-----PITSPP--CH 837
 QY 1083 TGPFTFVVPFLSSINGVPTTAATAAPVATSSPPNDISTSVIQSEVYVTEBGIAGVAT 1142
 DB 838 PPSF--PSPPISSQVSNPSPHPTSSPLPSSSTPR----- 871
 QY 1143 STGVVTSGLPLTPVESPVLSVSVSITIPAVVSISSPSSQVPTSEIYVSATLY 1202
 DB 872 -----FVPLPSQCP-----MSLPTTSPPTSP--CSQVTLSSPFR- 906
 QY 1203 PSVTYSATSAAGGTATGPKRPAPVAVSQAAAGSTTGATLTSSTTSPSTASQSLQ 1262
 DB 907 -----PPC-----PSTSEFPSTTAAPLS 925
 QY 1263 LSSSTSTPLAETVAVVSAHSLKTSHTSTGLAFSLASAPSSSSSPGAVSSYISQGLH 1322
 DB 926 LSS-----AFSLA----- 933
 QY 1323 PLVPSVIASTPLPQAAGPTSTPLLQVPSIPPLQVAVANYPVQGLIHSGQPALIP 1382
 DB 934 -----VMTAAQSL--SPSGGLIS 950
 QY 1383 NOPHTHCEVSDTQPKAGIDDIKTLBEKRLSPSEHSSGAQHASVLSLETSLVIESTV 1442
 DB 951 QSP----- 953

QY 1443 TPGIPTTAAPSKLSTSTSTCLPTNLPLGTVALPVTPTGVSTVSTTSGVKEG 1502
 Db 954 -----PA-----PPSLP-----SLPLPPVAFGGESP-SPTIAVE-- 985
 QY 1503 TABSKPPLTKAPVLPVGTLPAGTLPSEQLPPFPFGSLQSOQLDLDAQLARTLSPEK 1562
 Db 986 -SEASPP-----PARPLPGE----- 999
 QY 1563 ITVTSVAVGVSMAPTAITEAGTQPOKGVSYVEGVLATSSGAGVFKMGFPQVSYAADG 1622
 Db 1000 ----ALAIPLS-----EKGKPOL-----VGRFVYT----- 1020
 QY 1623 AOKGKXKEDAKSVAFESSTSESVLSSSPSTLVKEPENGITTPGISDVPESAHKT 1682
 Db 1021 ----SSKEPAPPLPLQPT--SPTLSSG-----PKPSTPQTLSSSDTEDEA--- 1060
 QY 1683 TMSKASDTGQPPKVRFPVOTTTANKVGRFSVSKTDDKITDTYKEGPVAPSPMDLEQAV 1742
 Db 1061 ----GGGP-----ETRELAESDR----- 1075
 QY 1743 LPAVTPKKEKPELSEPSHLNGPESDPEAFLSKVDVDGSGSPH---SPHOLSGSLPSON 1799
 Db 1076 ----AABGIGACVBEEDGDK-EPYVGSGSPQPLSHPSVWMN 1112
 QY 1800 LSGSLNSFNSSYMSDNESEDIEDLKLRLRLDKHLKEIQDLSROKHEIESLYTKL 1859
 Db 1113 VSYV-----SLCLSSSESESGDEDFWAFELQSLRQKHESEVETLQTLQKKEIEDLYSRL 1167
 QY 1860 GKTPPAVILPPAPPLSGRRRPFKSGS-KSSRSSSLGKNSPOLSG-----NLSSGSAAS 1913
 Db 1168 GKQPPGCIYAPAPMLSSRQR--LSKGSFPSTSRNSLQNSEPPGPGIMRNSLSGSGTGS 1225

RESULT 10

Q810H4 PRELIMINARY; PRT; 1222 AA.
 AC Q810H4;
 DT 01-JUN-2003 (TEMBLrel. 24, Created)
 DT 01-JUN-2003 (TEMBLrel. 24, Last sequence update)
 DT 01-OCT-2003 (TEMBLrel. 25, Last annotation update)
 DE MNK4 Ser/Thr Kinase.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MKY/NHed; TISSUE=Kidney;
 RA Toka H.R., Kahle K.T., Wilson F.H., Nelson-Williams C., Lifton R.P.;
 RL "Role of MNK4 and MNK1 in hypertensive rat strains."
 RT Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL, AY192567; AAC8858.1; -
 DR GO; GO:000524; F:ATP binding; IEA.
 DR GO; GO:0004674; F:Protein serine/threonine kinase activity; IEA.
 DR GO; GO:0004713; F:Protein-tyrosine kinase activity; IEA.
 DR GO; GO:0006468; P:Protein amino acid phosphorylation; IEA.
 DR InterPro: IPR000719; Prot. kinase.
 DR InterPro: IPR002290; Ser_Thr_kinase.
 DR InterPro: IPR006271; Ser_Thr_kinase.
 DR InterPro: IPR001245; Tyr_kinase.
 DR Pfam: PF00069; Pkinase; 1.
 DR ProDom: PD000001; Prot. kinase; 1.
 DR SMART: SM00220; S_TKc; 1.
 DR SMART: SM00219; T_YKc; 1.
 DR PROSITE: PS00011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
 KM Kinase.
 SQ SEQUENCE 1222 AA; 132832 MW; E753DAC9D1FA446 CRC64;

Query Match 19.3%; Score 2083; DB 11; Length 1222;
 Best Local Similarity 32.0%; Pred. No. 1.2e-92;
 Matches 598; Conservative 167; Mismatches 356; Indels 750; Gaps 53;

QY 103 POPSPDAVPGAPPEPHREETVTAATATQVAOQPPAAAP-----GEOAVAGPAP 153
 Db 24 PPSPTSMGPTLGPPEPRVRRFSGKABERPPSSRRSSVDLGLSSWSQPPASLLPBP 83
 QY 154 STVPSTSKDRVQPSLVGSKPEP-----PPASGSGGSAKPEP--- 193
 Db 84 PDPSPSAGPWR--SPSS--NSKEHPGWTGAAPVKAADSACELTVSSGGPQSRBPBV 139
 QY 194 ----QEERSQOQDDIEETKAVGMSNDGRFLKFDIEIGRGSFKTYVKGLDTEETVEAV 249
 Db 140 PDAARERRRRQBEKEDTETQAVATSPDGRVYKFDIEIGRGSFKTYVKGLDTEETVEAV 199
 QY 250 CELQDKLTYSRRQRFKEAEMLKGLQHPNIVRFYDSWETVKKKKCIYLVTELMTSGTL 309
 Db 200 CELQTKLSRAERQRFSEVEMKGLQHPNIVRFYDSWETVKKKKCIYLVTELMTSGTL 259
 QY 310 KYTKRFKMYKTKVLRSMCRQTLKGLQFHTTPPIIHDIDCDNIETFGPGSVYKGL 369
 Db 260 KYTKRFKMYKTKVLRSMCRQTLKGLQFHTTPPIIHDIDCDNIETFGPGSVYKGL 319
 QY 370 GLATLKRAFAKSVIGTPPFMAPEMYEKKYDSVDVYAFAGMGLMATSEYPSSECONNA 429
 Db 320 GLATLKRAFAKSVIGTPPFMAPEMYEKKYDSVDVYAFAGMGLMATSEYPSSECONNA 379
 QY 430 QYRRVTSQVGPASPDKVALPEVKEITEGCIPONKDERYSITDNLNHAFFQSETYRVRL 489
 Db 380 QYRRVTSQVGPASPDKVALPEVKEITEGCIPONKDERYSITDNLNHAFFQSETYRVRL 439
 QY 490 AEEDDEKTAIKLMLIEDIKLKGKYKDNMAIEPSFDLEPDPEPVAQEMVSGVYCG 549
 Db 440 AEEDDEKTAIKLMLIEDIKLKGKYKDNMAIEPSFDLEPDPEPVAQEMVSGVYCG 498
 QY 550 DHTKMAIKDRVSLIKRRRQROLVREOEKKKQESSILKQVESSASQTSIGIKOLPSA 609
 Db 499 DYQVAVARVRRRAALQRKREKLKARE-----LEVLPPD 533
 QY 610 STGIP-TASTTSASVT-QVEPEPADQHOQLQYQPSI-SYLSQTYDSSGGSVTFE 666
 Db 534 SGPPATVSMTPGPAPPAPEPEPADQHOQLQYQPSI-SYLSQTYDSSGGSVTFE 593
 QY 667 SRVSSQGYTSYSGHQALSTGTVPHIPISTVQAQSPHGVPPSSVAGQSGQGSSSS 726
 Db 594 ASDPALQP-----PGMPS--SPAEPLCLPSG----- 619
 QY 727 LTGVSSSPPIQHQOQGGIQQTPAPQQTQVYSLSQTSSEATTAQPVGQAPQVLPQV 786
 Db 620 -----FALS----- 627
 QY 787 SAGKSTQGVSVQAPAEVAVAPQATQPTTLASVDSASHDVASGMSDGENENVPSSGR 846
 Db 628 GPGSDPSPG-----DSYASDAASGLSD-----M 650
 QY 847 HEGRTTKHRYKRSVRSRSHKTSRPKRLIYNSKNGDRVEQLTNRKMYTFKFDLD 906
 Db 651 GEGQMKKPKVKTLLRR-----PKRLIYNSKNGDRVEQLTNRKMYTFKFDLD 904
 QY 907 GDNPEEIIATMVNNDRIILIERESFYDQVEIIEKADMLSEDSVSEPEGQGLIESLOGK 966
 Db 705 GDSPEEIIAAMVYNEFIIPERDGLSRIRREIQRVETLLKRD----- 747
 QY 967 DDYGFSGSQKLGEFQKQPIPASSMPQOIGIPTSLLQVHASARFIVSVPEBSRLRESK 1026
 Db 748 -----AG----- 749
 QY 1027 VFPEETDYVAASTAQSPGNLSHSAISLQAFSELRRAQTEGNTAPPMFSGTPT 1086
 Db 750 --PSEATEDALSPQEEPAAPALPGPSDAELQSSISPEQNS----- 788
 QY 1087 FVVPPEPLSSIAQVPTTAATAVAPATSPSPNDISTVQSEVTVPTBEGIAVATSTGV 1146
 Db 789 -----MAAFSTP-----SSP-----GTPLSPGT 807
 QY 1147 VTSGGLPIPVSSSPVLSVSVSSITTPAVVSIITSPSLQVPTSTSEIVVSTALPFSVT 1206

```

Db 808 PFSFG--TPVFPCPI-----FRTSSGCHP-----YFSSQ 836
QY 1207 VSASASAGSGTATPFGKRPAPVAVQQAAGSTTVATLTSTTSSPFTASQLSIQLSSS 1266
Db 837 VS-----SNPCQAPSSLLPSSSGAGV-----PFTS-----863
QY 1267 TSTPTLAETVVASASLIDKTSSTGTGLAFSLAPSSSSPGAGVSSYISQPGSLHPLVI 1326
Db 864 -----863
QY 1327 PSVASTPPIPLPQAAGPTSTPLLPQVPSIPLPQVAVNAVQQTLLHSQPPALLPQPH 1386
Db 864 PSL-----PTSSP-LPSPSPSY-----QVPLH-----PALLP-----890
QY 1387 THCEVSDSDTPKAPAGIDDKITLLEKARSLFSEHSSSGAGASVLSLTVIESVTPTGI 1446
Db 891 --CPS-----897
QY 1447 PTTAVAPSKLLTSTSTCLPPTMLPLGTVALPVPVTPGVSTPSTTSSGVKGTAPS 1506
Db 898 PSTTAAPLLSLASAFS-----LAVMTVA-----QSLISP-----SPGLLSQ 933
QY 1507 KPPLTKAPVLPVGTGLPAGTLPSEQLPPFPQPSLTGSOQPLEDLAQLRRTLSPEXITVT 1566
Db 934 SPP-----APPGPLPSMPL-----PLASCDQE-----955
QY 1567 SAQVPSMAAPTAIT--EAGTQPKGVSVQKEPVLATSSGAGVFKKGRQVSVADGAQ 1624
Db 956 -----SLAQIATEENEARPNQ-----PLGADARLAPI-----985
QY 1625 KEGKNSSEDAKSVFESSTSESSSVLSSSPSTLVKPEPNTITPGISDPVESAHTTA 1684
Db 986 -----985
QY 1685 SEAKSTGQPTKYGRFOVTTANKVGRFSVKTEKIDTDTKKEGPVAPSPFMDLEQAVLP 1744
Db 986 -----SEEGKQVLGRFQVTS-----KE--PAEPPLQASPTLSR 1019
QY 1745 AVIPKKEKPEL-----SEPSHLNGP-----SDPEAFSLRDVDDGSGSHSPHQLS 1791
Db 1020 SL--KLPPTQLTSESDTDESAAGPPTREALAESRAAGLQVAILDEEDKEKEPQIG 1077
QY 1792 SKSLPQNTLSQSLNSFNSSSYMS--DNESDIEDEDLKLRLRDLKHLKIDLOLSROKH 1850
Db 1078 SSPILSQPSVMMNYVSSLCSSSESGDEBEFAELQNLROKHLSEVALQTLQKK 1137
QY 1851 EISLTYTKLGKVPVAVIIPPAAPLSGRRRPTYSKGS--KSSRSSSLGNKSPOLSG-----1904
Db 1138 EIEDIVSRIGKOPPGIVAPAAMLSCORR--LSKGSFPTSRNSL--QESDLPGQIMR 1193
QY 1905 --NLGOSAS 1913
Db 1194 RNLSGSSSTGS 1204

```

RESULT 11

```

O811R5 ID O811R5 PRELIMINARY; PRT; 1210 AA.
AC O811R5:
DT 01-JUN-2003 (TEMBLrel. 24, Created)
DT 01-JUN-2003 (TEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
DE PRKNMKA.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=WKY; TISSUE=Kidney;
RA Zim Dahl H., Monti J., Hubner N.;

```

```

RT "Sequence analysis of WKNA in spontaneously hypertensive rats.";
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY187039; AAO18238.1; -.
DR PIR; A40135; A40135.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR GO; GO:0004713; F:protein tyrosine kinase activity; IEA.
DR GO; GO:0004648; F:protein amino acid phosphorylation; IEA.
DR InterPro; IPR00719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_kinase.
DR InterPro; IPR008271; Ser_thr_kinase.
DR InterPro; IPR01245; Tyr_kinase.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TKc; 1.
DR SMART; SM00219; TYKc; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW Hypothetical protein.
SQ
SEQUENCE 1210 AA; 131499 MW; 1C98471497DEC65B CRC64;

```

```

Query Match 19.3%; Score 2082; DB 11; Length 1210;
Best Local Similarity 31.8%; Pred. No. 1.3e-93;
Matches 593; Conservative 168; Mismatches 364; Indels 742; Gaps 50;

```

```

QY 103 PQSIRAPVQSPAPPRPHRETYTATATSGVAAQPPAAP-----GEQAVAGAP 153
Db 12 PPSLTSMGPTLGRPPRRRFRPSRAEPRFRSSRSRSDVDLGLSSWQPSALLPEP 71
QY 154 STPSSSTSKDRPVQSOGLVSKPEP-----PPANSGGSGSAXEP---193
Db 72 PDPDDAGPWR--SPSS--NSKEHPGWTGAAPVYAAASACELTVSSGPGSREPRV 127
QY 194 ----QEERSQQDDIELETKAVGMSNDGRFLKFDIEIGSFRTYKGLDTEETVEAV 249
Db 128 PDAARERRRQEKEDTQAVATSDDGRYLFKFDIEIGSFRTYKGLDTEETVEAV 187
QY 250 CELQDRKLTKSEQRKEPEMKLGLQHPIYVFTYSWSTYVKKCYIVLTBLMTSGTL 309
Db 188 CELQARKLTKSEQRKEPEMKLGLQHPIYVFTYSWSTYVKKCYIVLTBLMTSGTL 247
QY 310 KTYLRFKTYMKLVKLSWCKQLIKGLQFLHTRPPTIHHDKDNDLFIPTGSGSVTGL 369
Db 248 KTYLRFKTYMKLVKLSWCKQLIKGLQFLHTRPPTIHHDKDNDLFIPTGSGSVTGL 307
QY 370 GLATLRASFAKSVIGTPEFMAPEMYEKYDESDVYAFGCMLEAVATSEYPSKQMA 429
Db 308 GLATLRASFAKSVIGTPEFMAPEMYEKYDESDVYAFGCMLEAVATSEYPSKQMA 367
QY 430 QIYRYTSGVKASPPKVAIPYKEIIEGCIQNKQKERSIKDLNHAFFQETGYRVEL 489
Db 368 QIYRYTSGVKASPPKVAIPYKEIIEGCIQNKQKERSIKDLNHAFFQETGYRVEL 427
QY 490 AEDDGEKIAIKLMLRIEDIKLKGKYKNEAIEFSFDELRDVPEDVAQMEVSGYVCE 549
Db 428 AEDDGEKPGKLMMLMEDAR--GGRPRNQAIIEFLFQGRDAEVEADAMVALGLVCA 486
QY 487 DYQPARAVRERVAAIQKREKILKARE-----LEVLPPD 521
QY 610 STGIP--TASTTASVST--QVEPEPEADQHQLOVQOQPSI--SVLSNGTVDSGQSSVFPE 666
Db 522 GPPPATVYSMTPGPPSAPPEPEPEADQHQFLFEHAYVSTTSDBCTDGYLSSGFUD 581
QY 667 SRVSSQQTIVSYGSOHEQASHSTGVGHIPSTVQAOQOPGVPPSSVAGQSQGQGPSSSS 726
Db 582 ASDPMLP-----FGGMS--SAERHLCLPSC-----607
QY 727 LTGVSSQPTIQHPQQQGGIQTAPPOQTVOYLSQTSSTSEATTQAQVSPQAPQVLPQV 786
Db 608 -----FALSS-----IPRS 615

```

QY 787 SAGKOSTQGVSOVAPAEFPAVAPQATQPTTLASSVDSASGVASGMGDNENVPSSGR 846
 Db 616 GPGSDPSFG-----DYSASDAASGLSD-----M 638
 QY 847 HEGRTTKHRYKRSVRSRSHKTRRPKRLINYSNGDRVVEGQLTNHRKMTPEKFD 906
 Db 639 GBGQKQKQVKTLLRR-----PSSRLRVTSDSDSDRVNQCQLOTHNSKMTPEKFD 692
 QY 907 GDNPEIATIMVNDFTLAIERESFVDQVREITEKADEMLSQVSVPEGQGLLESLQCK 966
 Db 693 GDSPEIILAMVNEFLPSERDGLSRILEIQRVETLLKR----- 735
 QY 967 DDYFGSGQKLEGEFKQPIIPASSMPOQIGIPSSLTQVNHASGRFIVSPVESRLRESK 1026
 Db 736 -----AG----- 737
 QY 1027 VPPSEITDVAASTASPGMNLSSHASLSLQQAQFELRAQGTGEPNTAPNPSHTGPT 1086
 Db 738 --PSEATEDALSTQEEAAMPALPGPSDELORSISPEORS----- 776
 QY 1087 FPAVPPFLSSIAQVPTTAATAATAPVPAATSSPPNDISTSVQISEVTPTEEGIAGVATSTGV 1146
 Db 777 -----WAAFTSR-----SSP-----GTPLSPGT 795
 QY 1147 VTSGGPIPPVSESPVLSSVSSITTPAVVSISTTSPSLQVPTSTSEIVSSTALYPSVT 1206
 Db 796 PEPSPG--TPFPVPCPI-----FPITSPSCHP-----YPFQ 824
 QY 1207 VSAITSAAGSSTTPGKPPAVVSOQAAGSTYGALITVSTTTSPPTASQSLQSS 1266
 Db 825 VS-----SNPCQAPSSILPSSSGASQV-----PPPS----- 851
 QY 1267 TSTPTLAETVVAHSLDKTSHSSTTGLAFSLAPSSSSPGAGVSVYSIQPGCLAPLYI 1326
 Db 852 ----- 851
 QY 1327 PSYIATSTPLPOAGTSTPLPLPQVPSIPPLVQPVANVAVOQTLIHSPQAPALLPNOQH 1386
 Db 852 PSL-----PTSSP-LPFPSPSY-----QVPLH-----PASLPT-- 878
 QY 1387 THPEVDSDTQAPAGIDIKILEKRLSFHSHSSGQAHASVSLFSLIVESTYTPGI 1446
 Db 879 --CPS-----PPPL 885
 QY 1447 PTTAVAPSKLNTSTSTCLPPTNLPLGTVALPVTVPVTPQGVSTPSTTSGVKPETAAS 1506
 Db 886 PSTTAPPLSLASAFS-----LAWMTA--QSLSP-----SPGLSQ 921
 QY 1507 KPPLTKAPVLPGVTELPAGTLPSEQLPPPPPSLTQSQQPLBDLDAQLRRTLSPEXITVT 1566
 Db 922 SPF-----APGPLPSMPL-----PLASCQGE----- 943
 QY 1567 SANGVPSMAPTAIT--EAGTQOKGVSOYKEGVPULATSSGAGFFKMGFPQSVADGAQ 1624
 Db 944 -----SLSAQTETENASRNPAQ-----PLGDARLAPI----- 973
 QY 1625 KEGKNSKEDAKSVFESSSTSSSVLSSSPSTLVKPEPQGITIPGISDVPESAHTYA 1684
 Db 974 ----- 973
 QY 1685 SEAKSDTQOPTKVGRFQVTTANKVGRFVSXTEDKITDTKKGPAVAPPFMDLEQAVLP 1744
 Db 974 -----SSEKQQLVGRFOVTS-----KE--PAEPPLQASPFLSR 1007
 QY 1745 AVLPKKEKEL-----SEPSHNGP-----SSPEAFLSRVDVDSGSHSPQLS 1791
 Db 1008 SL--KLPTQLTSESSDTESSAAGPETREALAESDRAAGLVAIDEKDEKEKQIG 1065
 QY 1792 SKSLPQNLQSLSNFSNYSYMS--DNESDIEDEDLKLRLRDLKRLKEIIDLQSRQK 1850
 Db 1066 SPSILQPSVWMNYSSSLCLSSSESSESGEBEFMAHQLNQLKHLSEVALQTLQCK 1125
 QY 1851 EIESLYTKLGKVPAAVIPPADLSGRRRRPTKSKSSRSSSLGNKSP---QLSGNL 1906

Db 1126 EIEDLYSLGKQPPGIVAPAAMLSCORRLSKGSFPTSRNSLQSDLPFGIMKNSL 1185
 QY 1907 SGQSAAS 1913
 Db 1186 SSGSTGS 1192
 RESULT 12
 080UE6
 ID 080UE6 PRELIMINARY; PRT: 1222 AA.
 AC 080UE6;
 DT 01-JUN-2003 (TREMBLrel. 24, Created)
 DT 01-JUN-2003 (TREMBLrel. 24, Last sequence update)
 DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
 DE WNK4.
 GN PRKNK4.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 OX NCBI_TaxId=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6; TISSUE=Kidney;
 RX MEDLINE=22426867; PubMed=12515852;
 RA Wilson F.H., Kahle K.T., Sabath E., Lallio M., Rapson A.K.,
 RT Hoover R.S., Hebert S.C., Gamba G., Lifton R.P.;
 RT "Molecular pathogenesis of inherited hypertension with hyperkalemia:
 RT The Na-Cl cotransporter is inhibited by wild-type but not mutant
 RT WNK4";
 RL Proc. Natl. Acad. Sci. U.S.A. 100:680-684 (2003).
 DR EMBL: AY187027; AAC21955.1; -
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:004674; F:protein serine/threonine kinase activity; IEA.
 DR GO; GO:004713; F:protein-tyrosine kinase activity; IEA.
 DR GO; GO:006468; F:protein amino acid phosphorylation; IEA.
 DR InterPro: IPR000719; Prot Kinase.
 DR InterPro: IPR002290; Ser thr kinase.
 DR InterPro: IPR008271; Ser thr pkin_AS.
 DR InterPro: IPR001245; Tyr-kinase.
 DR Pfam: PF00069; pkinase; 1.
 DR ProDom: PD000001; Prot Kinase; 1.
 DR SMART: SM00220; S_TKC; 1.
 DR SMART: SM00219; TyrKc; 1.
 DR PROSITE: PS00101; PROTEIN KINASE DOM; 1.
 DR PROSITE: PS00108; PROTEIN KINASE ST; 1.
 SQ SEQUENCE 1222 AA; 132409 MW; BE0A193E7D79DCBE CRC64;
 Query Match 19.1%; Score 2068; DB 11; Length 1222;
 Best Local Similarity 31.8%; Pred. No. 6, 6e-93;
 Matches 601; Conservative 167; Mismatches 352; Indels 772; Gaps 55;
 QY 97 GRLS-----LQPSIPAAVPSGAPPEPHREIVTATASQVAAQPPPAAP----- 143
 Db 10 GVPMSQTEADLALRBPALTSSTGPTRLGPPRRVRRFSKABPRSSRSRSDVLGL 69
 QY 144 -----GEOAVAGAPASTVPSTSKDRPVSGPSLVGSKKEPP-----PAR 182
 Db 70 LSSMSQPSALLPBPDPDSAGPTR--SPPS--SKPEPBGTMGAAPVXAVDSACPEL 124
 QY 183 SSGSGG-SAKEP-----QEERSQQODDIELEIETAVAGMSNDGRELFKQIEIGRSFKT 234
 Db 125 TSSSGGPGSGREPPRPVDAARERRRRQKEKEDTQAVATSPDGRYLKFDIEIGRSFKT 184
 QY 235 VYKGLDTEITTVANWELQDRKLTKEBQRFKEBAEMLGLQHPNIVREFYDSNESTVXK 294
 Db 185 VYKGLDTEITTVANWELQDRKLTKEBQRFKEBAEMLGLQHPNIVREFYDSNESTVXK 294
 QY 295 KCIIVLTETMTSGTLTKYTKRKPKVMKIVLRSWCROILKGLQPLHTRPPIIHRDLKCN 354
 Db 245 VCIIVLTETMTSGTLTKYTKRKPKVMKIVLRSWCROILKGLQPLHTRPPIIHRDLKCN 304
 QY 355 IFITGFTGSKIGDGLATLTKASFAKSVIIGPEFMAPEMYEEKYDESVDVYAFGCMLE 414

```

Db      305 VFTTSGSGSVKIDLGATLKRAFSKAVIGTPEFMAPEMEYEKDEYDAVVAFGMCML 364
      415 MATSEPYSCQNAQAQYRVTSGVKAAPDXAIPRVKIIIEGCTRONDEXSTKDL 474
      365 MATSEPYSCQNAQAQYRVTSGVKAAPDXAIPRVKIIIEGCTRONDEXSTKDL 424
      475 NHAFFOETVAVELAEEDDEGEKIAIKLPIEDIKLKGKYNONEAIEFSFDLERDVE 534
      425 AAHAFRERGVHLEAEDDEGEKGLKLMEDAR--GGRPDNOAIEFLFOLGRDAE 483
      535 DVAQENWESGYVEBGDKTMAKAIKORVSLIKRREGQIVREOEKKEQESSLKQVE 594
      484 EVAQEWALGLVEADYQVAVARERVAAIORREKLRKARE----- 526
      595 QSSASQGTGIKQLSASTGIPTASTTASVSTQV---EPEEPADQOQLOLOQOPST-SVL 650
      527 -----LEVLPPDS--GPPPATVSLAPGPPSAFPEPEPEEADQHQSLFRHASYSTT 577
      651 SDGTVDGQSSSVFTESRVSQOQTVSYGQHEQAHSTGVGHIPSTVQAQSPHGYVP 710
      578 SDGETDGLSSSGFLASDPALQ-----PGLPSS----- 608
      711 SSVAGQSGQSGPSSSITGVSSSQPIQHPQOQCGIQQTAPQQTVOYSLQSTTSSEATT 770
      609 -----PAESHLCIPSG-----FALS----- 623
      771 AQVSPQAPQVLPQVSAGKQSTQGVQVAPAPVAVAPQATQPTTLASVDSASDVA 830
      624 -----LPRGPGSDPSPG-----DSVASDA 644
      831 SGMSDGENVPSSSGHEGRTTKRHKYSVRSRHEKTSRPKILILNYSNKGDRVCEQ 890
      645 SGLSD-----MGEQGQMKNPVKTLRR-----PRSLRTVTSVSDQDRVCEQ 688
      891 LEHNRKMTFKEDLDGDNBEIATIMWNNDFILABESFVDQVREIEKADMLSEVY 950
      689 LQTHNSKMTFRDLQDSDPEEIAAMVYNEFILPSRDFLSRIEIIQIRVETLLKRD- 747
      951 SVPEBDQGLSEIQGKDDYGFSGSQKLEGEFKQPIASHPQOIGIPTSLTLQVHSAKR 1010
      748 -----Ag- 749
      1011 RPIVSPVESRLRESKVFSEITDTVAASTAGSPGMNLSHASLSLQOAFSELRAQMT 1070
      750 -----PPE-----AABALSP-----Q 761
      1071 EGPNTAPPNFSHTGTFPVVPPFLSIAGVPTTAAATAPVATSPSPNDISTSVIQSEVT 1130
      762 EEPALP-----ALPQPMAPQORSIS----- 783
      1131 VPEBEGIAGVATSTGVVSGGLPIPEVSSPVLSSVSSITIPAVVISITSLQVPTS 1190
      784 -PEQRMAAFSTSP---SSEGTPLSP----- 805
      1191 TSEIIVSSTALVSVTASATASAGSGTAPGKPPAVVSOQAAGSTVGATLTLSVSTT 1250
      806 -----CAPSPGPP----- 815
      1251 SEPSTASQSLQSLSSSTPTLAETVAVSAHSIDKTSHTSSTGLAFSLAPSSSSPPGAG 1310
      816 VPECP-----IPPIPSPCYCPGPFQ 836
      1311 VSS--YISQGGHLPLVIPSVLASTPILPQAAGPTSTPLLPOVPSLP--PLVQPVANVA 1366
      837 VSSNPYPQAPSSILPL--SSASQVPLPSSSLPISAP--LPSSVSPQDPL----- 883
      1367 VQCTLHSPQAPALLPNOPHTHCPVDSDTQPKAPGIDIKTLEKRLSLFSEHSSGQAQ 1426
      884 -----SPISLP-----VCPSS----- 893
      1427 HASVSLSLVLESTVTPGIPPTTAAVPSKLLTSTSTCLPPTNLPLGVALLPVTVPVTPG 1486

```

```

Db      894 -----PSPSTTAAPLSLASAFS-----LAVTVA---QSLTSP- 926
      1487 QVSTPSTTSSGVKPPGTAASKPPLTAPVLVGTETLPACTLPBEOQLPPFPGPSLTOSQOP 1546
      927 -----SPGILSQSP-----APPGPLPSL----- 945
      1547 LEDLDQRLRTTSLPEXITVTSVAVPVSMAAPTATENGTPQKGVSGVKGVPILATSSGA 1606
      946 -----PLSLASCD----- 953
      1607 GYFKMGRFQVSVAAQAKGKNSDASKSVHESSTSESSVLSGSPESSTLVKPEPENG 1666
      954 -----GESLSAQAEENE-----ASNPAPPL----- 977
      1667 TIRGISDVPEASHKTTASAK-----SDTGQPTKVGRFQVTTANAKVGRSVKTEKXT 1722
      978 -----GDARLAPISEBKPLVORFQVTS----- 1002
      1723 DTKKEGPAVSPPMDEQAV-----LPAVLPKKEKELSPSHLNG-----SSDPEA 1770
      1003 ---KE--PAEPPLQAPSPILSRSLKLPSP-PLTSSSDTSDSAGGPETREALABDRAA 1056
      1771 AFLSRVDGSGSPHSPHOLSSKSLPSQNLSSLSNFSNYSYS--DNESDIEDDLKTE 1829
      1057 EGLGVAVDEKDEGKPLLGSSSPILSHSPVWMNYSYSLCISSESSSGEDEEFMAE 1116
      1830 LRRLRDKHLKEIODLOSQKHETESLYTKLGKPPRAVILPPAPPLGRRRRPRTKSGS-K 1888
      1117 LQMLRQKHLSEVALQTLQKKEILEDYLSRLGQPPGIVAPAMLSRCORR--LSKGSPP 1174
      1889 SSRSSSLGNKSPOLSG-----NLGQSAAS 1913
      1175 TSTRNLT--QRGDLPPQGITMRNRSLSGSSSTGS 1204

Db      1175 TSTRNLT--QRGDLPPQGITMRNRSLSGSSSTGS 1204

RESULT 13
Q8N673
ID Q8N673 PRELIMINARY; PRT; 393 AA.
AC Q8N673;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OC NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RC TISSUE=Placenta;
RU Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC021121; AAH21121.1; -.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006468; F:protein amino acid phosphorylation; IEA.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_kinase.
DR InterPro; IPR008271; Ser_thr_kinase.
DR InterPro; IPR001245; Tyr_kinase.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TKc; 1.
DR SMART; SM00219; TyKc; 1.
DR PROSITE; PS00011; PROTEIN KINASE DOM; 1.
DR PROSITE; PS01008; PROTEIN KINASE ST; 1.
DR Hypothetical protein: amp-binding; Transferase.
SQ
SEQUENCE 393 AA; 42434 MW; 89555C4EF15BFD3E CRC64;

```

Query Match 18.3%; Score 1974; DB 4; Length 393;
 Best Local Similarity 100.0%; Pred. No. 5.5e-89;

Matches 385; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSGGAAEKOSTSTGSLFLSPAPAPKNGSSSVGEKGAADAVTGRTEYRRRRAT 60
 1 MSGGAAEKOSTSTGSLFLSPAPAPKNGSSSVGEKGAADAVTGRTEYRRRRAT 60
 Db 1 MSGGAAEKOSTSTGSLFLSPAPAPKNGSSSVGEKGAADAVTGRTEYRRRRAT 60
 QY 61 MDKDSGAAATTTTTHRRFRFRSVICDSNATLPLPLSLPQSPAPAPQAPPEPH 120
 61 MDKDSGAAATTTTTHRRFRFRSVICDSNATLPLPLSLPQSPAPAPQAPPEPH 120
 Db 61 MDKDSGAAATTTTTHRRFRFRSVICDSNATLPLPLSLPQSPAPAPQAPPEPH 120
 QY 121 REETVATATSOVAQOPPAAPAAAGVAVGAPASTVPSSTCDRPPSPQSYLVSKKEPP 180
 121 REETVATATSOVAQOPPAAPAAAGVAVGAPASTVPSSTCDRPPSPQSYLVSKKEPP 180
 Db 121 REETVATATSOVAQOPPAAPAAAGVAVGAPASTVPSSTCDRPPSPQSYLVSKKEPP 180
 QY 181 AASGSGGSAKEPQERSSQODDIELETKAVGMSDGFLEKFDIEIGSFYTKYKGD 240
 181 AASGSGGSAKEPQERSSQODDIELETKAVGMSDGFLEKFDIEIGSFYTKYKGD 240
 Db 181 AASGSGGSAKEPQERSSQODDIELETKAVGMSDGFLEKFDIEIGSFYTKYKGD 240
 QY 241 TETTVAVANCELODRKLTSEKRFYEAEMLKGLQHPNIVRFYDSWESTVKKKCIIV 300
 241 TETTVAVANCELODRKLTSEKRFYEAEMLKGLQHPNIVRFYDSWESTVKKKCIIV 300
 Db 241 TETTVAVANCELODRKLTSEKRFYEAEMLKGLQHPNIVRFYDSWESTVKKKCIIV 300
 QY 301 TELMTSGITKTYLKRKWKIKYLRSCQILKGLQFLHTRPPIIHRDLKCDNFIITGP 360
 301 TELMTSGITKTYLKRKWKIKYLRSCQILKGLQFLHTRPPIIHRDLKCDNFIITGP 360
 Db 301 TELMTSGITKTYLKRKWKIKYLRSCQILKGLQFLHTRPPIIHRDLKCDNFIITGP 360
 QY 361 TGSVKIGDGLATLKRKSPAKSVYIG 385
 361 TGSVKIGDGLATLKRKSPAKSVYIG 385
 Db 361 TGSVKIGDGLATLKRKSPAKSVYIG 385

RESULT 14
 Q80XN2 PRELIMINARY; PRT: 1048 AA.

AC Q80XN2; 24, Created
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE 2010002J1IRLK protein (Fragment).
 GN GN 2010002J1IRLK
 OS Mus musculus (mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OC NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=FVB/N; TISSUE=Kidney;
 RA MEDLIN=22388257; PubMed=12477932;
 RA Klausner R.D., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Donald M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loggellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McMan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W.,
 RA Villalon D.K., Murthy D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahney J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimm J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywicki M.I., Skalska U., Smalins D.E., Schnerch A., Schein J.E.,
 RA Jones S.J., Maitra M.A.,
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RP SEQUENCE FROM N.A.
 RC STRAIN=FVB/N; TISSUE=Kidney;
 RA Klausner R.;
 RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.

DR EMBL: BC043677; AA043677.1; -.
 DR GO: GO:0005524; F:ATP binding; IEA.
 DR GO: GO:0004674; F:protein serine/threonine kinase activity; IEA.
 DR GO: GO:0004713; F:protein-tyrosine kinase activity; IEA.
 DR GO: GO:0004668; F:protein amino acid phosphorylation; IEA.
 DR InterPro: IPR000719; Prot Kinase.
 DR InterPro: IPR002290; Ser thr kinase.
 DR InterPro: IPR008271; Ser thr kinase.
 DR InterPro: IPR001245; Tyr kinase.
 DR Pfam: PF00069; pkinase; 1.
 DR ProDom: PD000001; Prot. kinase; 1.
 DR SMART: SM00220; STK; 1.
 DR SMART: SM00219; TyKc; 1.
 DR PROSITE: PS00011; PROTEIN KINASE DOM; 1.
 DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
 FT NON_TER 1
 SQ SEQUENCE 1048 AA; 113942 MW; 45F8BAAB6D18DA66 CRC64;

Query Match 17.7%; Score 1911.5; DB 11; Length 1048;
 Best Local Similarity 32.1%; Pred. No. 2.5e-85;
 Matches 551; Conservative 146; Mismatches 297; Indels 725; Gaps 48;

QY 228 GGSFRTYKGLDITTVAVANCELODRKLTSEKRFYEAEMLKGLQHPNIVRFYDSW 287
 228 GGSFRTYKGLDITTVAVANCELODRKLTSEKRFYEAEMLKGLQHPNIVRFYDSW 287
 Db 228 GGSFRTYKGLDITTVAVANCELODRKLTSEKRFYEAEMLKGLQHPNIVRFYDSW 287
 QY 4 GRGSEKTYVKGADITTVAVANCELODRKLTSEKRFYEAEMLKGLQHPNIVRFYDSW 63
 4 GRGSEKTYVKGADITTVAVANCELODRKLTSEKRFYEAEMLKGLQHPNIVRFYDSW 63
 Db 4 GRGSEKTYVKGADITTVAVANCELODRKLTSEKRFYEAEMLKGLQHPNIVRFYDSW 63
 QY 288 ESTVKGKCIIVATVTEMTSGITKTYLKRKWKIKYLRSCQILKGLQFLHTRPPII 347
 288 ESTVKGKCIIVATVTEMTSGITKTYLKRKWKIKYLRSCQILKGLQFLHTRPPII 347
 Db 288 ESTVKGKCIIVATVTEMTSGITKTYLKRKWKIKYLRSCQILKGLQFLHTRPPII 347
 QY 64 KSVLRGQVCIVATVTEMTSGITKTYLKRKWKIKYLRSCQILKGLQFLHTRPPII 123
 64 KSVLRGQVCIVATVTEMTSGITKTYLKRKWKIKYLRSCQILKGLQFLHTRPPII 123
 Db 64 KSVLRGQVCIVATVTEMTSGITKTYLKRKWKIKYLRSCQILKGLQFLHTRPPII 123
 QY 348 RDLKCDNFIITGPTSVKIGDGLATLKRKSPAKSVYIGTPEFADMEYEKYDESIVYA 407
 348 RDLKCDNFIITGPTSVKIGDGLATLKRKSPAKSVYIGTPEFADMEYEKYDESIVYA 407
 Db 348 RDLKCDNFIITGPTSVKIGDGLATLKRKSPAKSVYIGTPEFADMEYEKYDESIVYA 407
 QY 124 RDLKCDNFIITGPTSVKIGDGLATLKRKSPAKSVYIGTPEFADMEYEKYDESIVYA 183
 124 RDLKCDNFIITGPTSVKIGDGLATLKRKSPAKSVYIGTPEFADMEYEKYDESIVYA 183
 Db 124 RDLKCDNFIITGPTSVKIGDGLATLKRKSPAKSVYIGTPEFADMEYEKYDESIVYA 183
 QY 408 FQCMLEMAATSEYSECCNAAQIYRKYTSYKPKSPFQVVAIPYKELLGICRQNKDER 467
 408 FQCMLEMAATSEYSECCNAAQIYRKYTSYKPKSPFQVVAIPYKELLGICRQNKDER 467
 Db 408 FQCMLEMAATSEYSECCNAAQIYRKYTSYKPKSPFQVVAIPYKELLGICRQNKDER 467
 QY 184 FQCMLEMAATSEYSECCNAAQIYRKYTSYKPKSPFQVVAIPYKELLGICRQNKDER 243
 184 FQCMLEMAATSEYSECCNAAQIYRKYTSYKPKSPFQVVAIPYKELLGICRQNKDER 243
 Db 184 FQCMLEMAATSEYSECCNAAQIYRKYTSYKPKSPFQVVAIPYKELLGICRQNKDER 243
 QY 468 YSIKDLNNAFPOEFGVAVELAEEDDGEKIKLMLRREDIKKLGKXKNDKATFEFSD 527
 468 YSIKDLNNAFPOEFGVAVELAEEDDGEKIKLMLRREDIKKLGKXKNDKATFEFSD 527
 Db 468 YSIKDLNNAFPOEFGVAVELAEEDDGEKIKLMLRREDIKKLGKXKNDKATFEFSD 527
 QY 244 FTIQLLNAFPRERGVAVELAEEDDGEKIKLMLRREDIKKLGKXKNDKATFEFSD 302
 244 FTIQLLNAFPRERGVAVELAEEDDGEKIKLMLRREDIKKLGKXKNDKATFEFSD 302
 Db 244 FTIQLLNAFPRERGVAVELAEEDDGEKIKLMLRREDIKKLGKXKNDKATFEFSD 302
 QY 528 IERDVPEDVADQVSGGVCEGDHKTMAKIDRYSILKRKEQRLVREDEKKKQES 587
 528 IERDVPEDVADQVSGGVCEGDHKTMAKIDRYSILKRKEQRLVREDEKKKQES 587
 Db 528 IERDVPEDVADQVSGGVCEGDHKTMAKIDRYSILKRKEQRLVREDEKKKQES 587
 QY 303 IGRDAAEVAGQEMVALGLVCEADYQVAVARAVERVAIQRKEXLRKARE 352
 303 IGRDAAEVAGQEMVALGLVCEADYQVAVARAVERVAIQRKEXLRKARE 352
 Db 303 IGRDAAEVAGQEMVALGLVCEADYQVAVARAVERVAIQRKEXLRKARE 352
 QY 588 SLKQVQSSASQGTIKQLPSASTGIPTASTTSASVSYV 644
 588 SLKQVQSSASQGTIKQLPSASTGIPTASTTSASVSYV 644
 Db 588 SLKQVQSSASQGTIKQLPSASTGIPTASTTSASVSYV 644
 QY 353 -----LEVLPDPS-GPPATVSIAGPPAPPEPEPEPADHOGFL 396
 353 -----LEVLPDPS-GPPATVSIAGPPAPPEPEPEPADHOGFL 396
 Db 353 -----LEVLPDPS-GPPATVSIAGPPAPPEPEPEPADHOGFL 396
 QY 645 PSI-SVLSDGVTVDSGGSSVFVESRVSSQGTYSQSHQAHSTGTPGHIPSTVQAQSQ 703
 645 PSI-SVLSDGVTVDSGGSSVFVESRVSSQGTYSQSHQAHSTGTPGHIPSTVQAQSQ 703
 Db 645 PSI-SVLSDGVTVDSGGSSVFVESRVSSQGTYSQSHQAHSTGTPGHIPSTVQAQSQ 703
 QY 397 AYSSTSDCEFDGYSLSGFLDASDPAQP 434
 397 AYSSTSDCEFDGYSLSGFLDASDPAQP 434
 Db 397 AYSSTSDCEFDGYSLSGFLDASDPAQP 434
 QY 704 PHGVYPPSSVAGQSGQSPSSSLTGVSQPIQHPOQOQGIQQAAPQQTQVSLQTS 763
 704 PHGVYPPSSVAGQSGQSPSSSLTGVSQPIQHPOQOQGIQQAAPQQTQVSLQTS 763
 Db 704 PHGVYPPSSVAGQSGQSPSSSLTGVSQPIQHPOQOQGIQQAAPQQTQVSLQTS 763
 QY 435 -----PAESHLCPLSG-----FALS 449
 435 -----PAESHLCPLSG-----FALS 449
 Db 435 -----PAESHLCPLSG-----FALS 449
 QY 764 TSEATTAQPVQOPAPQVLPVQYSAKQSTQGVQVAPAPVAVQAPATOPTTLASVD 823
 764 TSEATTAQPVQOPAPQVLPVQYSAKQSTQGVQVAPAPVAVQAPATOPTTLASVD 823
 Db 764 TSEATTAQPVQOPAPQVLPVQYSAKQSTQGVQVAPAPVAVQAPATOPTTLASVD 823
 QY 450 -----IPRSGPSDPSG-----D 463
 450 -----IPRSGPSDPSG-----D 463
 Db 450 -----IPRSGPSDPSG-----D 463
 QY 824 SAHSDVASGMDGNENVPSSSRHGEKRTKRYKRSVRSRSHKTSPEKRLILNYSNG 883
 824 SAHSDVASGMDGNENVPSSSRHGEKRTKRYKRSVRSRSHKTSPEKRLILNYSNG 883
 Db 824 SAHSDVASGMDGNENVPSSSRHGEKRTKRYKRSVRSRSHKTSPEKRLILNYSNG 883
 QY 464 STASDASGSD-----MGEGCKRKAPVYTLKR-----PSRLKTVSDQS 507
 464 STASDASGSD-----MGEGCKRKAPVYTLKR-----PSRLKTVSDQS 507
 Db 464 STASDASGSD-----MGEGCKRKAPVYTLKR-----PSRLKTVSDQS 507
 QY 884 DRVVECOLETHNRKRVTFKFDLDGNPEEIIATIMVNDIFILARESFDVQVEIIEKAD 943
 884 DRVVECOLETHNRKRVTFKFDLDGNPEEIIATIMVNDIFILARESFDVQVEIIEKAD 943
 Db 884 DRVVECOLETHNRKRVTFKFDLDGNPEEIIATIMVNDIFILARESFDVQVEIIEKAD 943
 QY 508 DRVVECOLQTHNSKRVTFKFDLDGDSPEEIIAAMVNEFIIPSEBDGFLSRIEIIQRYE 567
 508 DRVVECOLQTHNSKRVTFKFDLDGDSPEEIIAAMVNEFIIPSEBDGFLSRIEIIQRYE 567
 Db 508 DRVVECOLQTHNSKRVTFKFDLDGDSPEEIIAAMVNEFIIPSEBDGFLSRIEIIQRYE 567
 QY 944 EMLSEVSVPEBGDGLSLQKQDYGRSGSKLGEFQKQIPASSMPOQIGLPTSLTG 1003
 944 EMLSEVSVPEBGDGLSLQKQDYGRSGSKLGEFQKQIPASSMPOQIGLPTSLTG 1003
 Db 944 EMLSEVSVPEBGDGLSLQKQDYGRSGSKLGEFQKQIPASSMPOQIGLPTSLTG 1003

```

QY 1004 VYHAGRRFVSPVPSRLREKVPSEITDVTAASTQSPGMNLSHSSASLSLQCAFSE 1063
Db 574 ---AG-----PPE-----AAEDALP----- 586
QY 1064 LRRRAOMTEGNTAPPNFSHTGPTFPVVPPLSSIAVPTTAAATAPVATSPENDISTS 1123
Db 587 ---QEEPAALP-----ALPPEPAAEQORIS-- 609
QY 1124 VIOSEVTPREBEGIAGVATSTGVVTSGLPIPVSESPVLSVSSVSTTTPAVVISITSTP 1183
Db 610 ---PQORWMAFSTSP---SSPPTLSP----- 631
QY 1184 SLQVPTSTSEIVVSTALYPSVTVAATSAAGSTATPGKPPAVVSOQAAGSTTVGATL 1243
Db 632 ---CAPSPGTPP----- 641
QY 1244 TSVSTTSPSTASQSLQSSSTPTLAEIVVSAHSLDKTSHSSTTGIAFSLAPSG 1303
Db 642 ---VFPCP-----IFPITSBSC 655
QY 1304 SSSPAGVSS--YISQPGHPIVPSVIASTPILQAGPTSTPLLPQVPSIP--PLYO 1359
Db 656 YPCFSPQVSSNPYPQAPSSILPL---SSSASQVPLPSSSLPIBAP-LPFSFSPYODPL-- 709
QY 1360 PVANVAVOQOTLIHSQPOPALLPNOPHTHCPEVDSPTQPKAPGIDDIKTLKELRSLFSE 1419
Db 710 ---SPTSLP---VCPSS----- 719
QY 1420 HSSSGAQAHSVLSLETSLVIESTVPGITPTTAVAPSKLSTSTSTCLPPTNLPLGTVALPV 1479
Db 720 ---PFSLPSTTAPPLSLIASAFS-----LAWMTVA-- 746
QY 1480 TPVYTPGCVSTPVSTTSGVKPPTASKPPLTAAPVLPVGTETPACTLPSBQLEPPPGCS 1539
Db 747 QSLTSP---SPGLSQSP-----APPGPLPS 770
QY 1540 LTQSQOPLDLDAQLRRLTSPKXITVTSAGVPYMAAPTAITEAGTQPOKGVQVKEGPV 1599
Db 771 L-----PLSLASCD----- 779
QY 1600 LATSSGAVPKMGKRFQVSVAAADQAQKGNKSEDAKSVAHESSTSSSVLSSSPSTLV 1659
Db 780 ---QESTLAQTAETENE-----ASNPAQPLL 803
QY 1660 KPBNNGITPGISSDVPESAHKTASEAK---SDTQCPKVGKRPQVTTANKVGFVS 1715
Db 804 ---GDARLAPISEBGPOLVGRFOVTS-- 828
QY 1716 KTEEDKITDKKEGPVAPSPFMDLEQAV---LPVAPIPKKEKPELSEPSHLNGP----- 1764
Db 829 ---KE--PAEPPLQAPASPTLSSKLTPSP-PLTSESDPTEDDAAGGFETREAL 875
QY 1765 -SSDPPAATFLSRVDVDDSGSPHSFOLSSKSLPSQNLSQLSNSFNSSYMS--DNESDIE 1822
Db 876 AASDRAAEELGVAVADDEKDEKGEKPLIGSSPILSHSPVMWYSSYSLCISSESSSGB 935
QY 1823 DEDLKLRLRLDKHKEIODLOSROKHELESYTLGKVPVAVITPPAAPLSGRRRT 1882
Db 936 DEEPWAELONLKOKHLEVEALQTLQKKEIEDLYSLGKOPPGIAPAAMLSCROAR-- 993
QY 1883 KSKGS-KSSRSSLSGNKSPOLSG-----NLSGQSAAS 1913
Db 994 LSKGSFPTSRNLSL-QRSDLPQPGIMRNKNSLSSSTGS 1030

```

```

DE AC2-059.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxId=10116;
RN [1]
RP SEQUENCE FROM N.A.
RA Xu C.S., Li W.Q., Li Y.C., Yang K.J., Yan H.M., Chang C.F., Zhao L.F.,
RA Ma H., Wang L., Wang S.F., Han H.P., Wang G.P., Chai L.Q., Yuan J.Y.,
RA Shi J.B., Rahman S., Wang Q.N., Zhang J.B.;
RT "Liver regeneration after PH.";
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY32328; AAP66260.1; -
SQ SEQUENCE 1552 AA; 173188 MW; 23AA592C8CDD088CA CRC64;

Query Match 17.5%; Score 1890.5; DB 11; Length 1552;
Best Local Similarity 42.7%; Pred. No. 4,6e-84;
Matches 452; Conservative 125; Mismatches 236; Indels 245; Gaps 28;

QY 103 PQSIPAAVPGAPRPPHREVTATATQVAAQPPAAAP-----GEQAVAGAP 153
Db 185 PPSPLTSMGPTRLGPPRRVRRPFGAEPSPRSRBSVDLGLSWSQPSLLEP 244
QY 154 STVPSSTSKDREVSQPSLGSKEP-----PPASGSGGSAKEP-- 193
Db 245 PDPPDSAGPMR--SPSS--NSKHEPBGTYTGAAPVAAADSACELTVSSGGPSRPPRV 300
QY 194 ---QERSQOQDDIELEETKAVGMSNGRFLKPDIEIGRGSFITYKGLDTEVAVM 249
Db 301 PDAAARERREDEBEKEDTQAVATSPDGRYLKFDIEIGRGSFITYKGLDTEVAVM 360
QY 250 CELQORUKLKSRRORPEKEAEMLKGLQHENVIRFYMSWSTYKGGKCYLVNLMMSGTL 309
Db 361 CELQTRKLSRARORPSESEVEMKGLQHPNIVRFYDMSVLRGQVCYLVNLMMSGTL 420
QY 310 KYLKKFKYMKIKVLRSMCRQILKGLQFLATRTPTPIIHRDLKDNFTIGTSPYKIGDL 369
Db 421 KYLRRFRMRKRVLRQWRQRLRGLHFLHSVPTLIRHDLKDNVFTIGSPSYKIGDL 480
QY 370 GLATLKRAFPASVIGTPEFPAPDEMYEKEYDSVDVYAFGMCLKEMATSEYPSBCONAA 429
Db 481 GLATLKRAFPASVIGTPEFPAPDEMYEKEYDAVDVYAFGMCLKEMATSEYPSBCONAA 540
QY 430 QYRRTSGVKKASPEKVAIPVKEIIEGCTRONDERISINDLNHAFQETGRVHEL 489
Db 541 QYRKVTSGTKNSTFYKVMPEVKELIBGCLRTDKNERFTIQDLTHAFRFRERGVHEL 600
QY 490 AEEDGEEKIATLMLRIEDIKLKQKYNKNEAIEFSFDELPDPVBDVAQEMWESGVYCEG 549
Db 601 AEEDGEEKGLKMLRMEDARR--GGRPRNQAIIEFLFQGRDAAEVAAQEMVALGLVCEA 659
QY 550 DKTMAKAIKDNVSLIKRRRQOLVREBOEKKKQOEBSSLKQOEVOSASQTSIGKOLPSA 609
Db 660 DYQPARAVRERVAALQREKELIRARE-----LEVLPPD 694
QY 610 STGIP-TASTSASVT-QVEPEPEADQHQLOQYQPSI-SVLSDGTVDGSGSSVFPE 666
Db 695 SCPPPATVSMTGPPSAPFPEPEPEADQHQFLFRHASYSTTSCEFDGYLSSGFLD 754
QY 667 SVSSSQGVTSYSGOHOAHSTGVPGHLPSTVQAQSQPHGVVPPSSVAQOGSGQPSSSS 726
Db 755 AADPALQ-----PGCMS---SPARBP-LCLPVFALISIRSGGS-- 792
QY 727 LTGVSSQPTIQHPQOQOQIGQTAPPOQTVQYLSQTSSTSEATTQAPVYQCPAPVYLPV 786
Db 793 ---DPSFGDMVYLVSVQAT-----PQWOPQ- 816
QY 787 SAGKSTGQGVQVAPAEVAVAPQATPTTLASSVDSHSDVYASGMSGNENWSSSGR 846
Db 817 ---ALSDMG----- 822
QY 847 HEGRTTKHYRKSVAHSRHEKTSRPKLRILNVSNKGDVYVCOLETHNRKAVTFEKDLD 906

```

RESULT 15
 Q7TPK6 ID PRELIMINARY; PRT; 1552 AA.
 AC Q7TPK6;
 DT 01-OCT-2003 (TRENBLrel. 25, Created)
 DT 01-OCT-2003 (TRENBLrel. 25, Last sequence update)
 DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)

```

Db      823 -EGGQKRNKPVKTLRR-----PRSRKRVTSVSDQSDRVBECQLOTHNSKMYTFRRDLD 875
QY      907 GDNPEETATIMVNDPILAIERESFYDQVREIIEKADML-----SEDVSVPEGD 957
      ||:||||| ||:||||| ||:||||| ||:||||| ||:|||||
Db      876 GDSPEETIAAMVYNEFILPSEKDFLSRIREIIQVETLTKRDAGPSEATEDALSPQEP 935
      ||:||||| ||:||||| ||:||||| ||:||||| ||:|||||
QY      958 QGLESLOGKDDYGFSGSQKLEGEF-----KQPI-----PAS-SMPQOIGIFTSSLTQVH 1006
      ||:||||| ||:||||| ||:||||| ||:||||| ||:|||||
Db      936 AAMPALPGPSD---EGKFOIVGRFOVTSKEPAEPPIQPAFTLSRLKLPQITS--E 990
      ||:||||| ||:||||| ||:||||| ||:||||| ||:|||||
QY      1007 SAGRRIIVSPVPSR--IRESKVFPSEITDVTAASTAQSPG----- 1045
      ||:||||| ||:||||| ||:||||| ||:||||| ||:|||||
Db      991 SSPTEDSAAGPFTREALAESDRABGL--GVAIDBEKDEKKEPQIGSSPILSQSPVW 1048
      ||:||||| ||:||||| ||:||||| ||:||||| ||:|||||
QY      1046 MNLSSHASLSLQOAFS-----ELRRQOMTE 1071
      ||:||||| ||:||||| ||:||||| ||:||||| ||:|||||
Db      1049 MNYYSLSLCSSESSESSEGEDEFMAELQNLKQKHLSF 1086
      ||:||||| ||:||||| ||:||||| ||:||||| ||:|||||

```

Search completed: September 24, 2004, 01:11:02
 Job time : 233 secs

Blank

Sheet

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: September 25, 2004, 01:25:13 ; Search time 11801 Seconds

(without alignments)
7845.159 Million cell updates/sec

Title: US-10-010-720-14

Perfect score: 10812
Sequence: 1 MSGGAARKOSTPESLFLSP.....NISNLQKISINPGSNLRIT 2136

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgmt1/USPTO.spool/US10010720/runat.23092004.154319.25242/app.query.fasta.1.2311
-DB=GenEmbl -OFMT=fastap -SUFFIX=rge -MINMATCH=0 -LOOFCU=0 -LOOFEVT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blomsum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=prco -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10010720 @CGM 1.1 8347 @runat.23092004.154319.25242 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -MAIN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl.*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_cm:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pac:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*

29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rdi:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	10810	100.0	6411	6 AR302594	AR302594 Sequence
2	10794.5	99.8	7280	6 AX766337	AX766337 Sequence
3	10753.5	99.5	6690	6 AR302591	AR302591 Sequence
4	10677	98.8	7149	6 AR302588	AR302588 Sequence
5	10677	98.8	7149	6 HSA296280	AJ296290 Homo sapi
6	10646	98.5	6327	6 AR302602	AR302602 Sequence
7	10589.5	97.9	6606	6 AR302599	AR302599 Sequence
8	10513	97.2	7065	6 AR302586	AR302586 Sequence
9	10508	97.2	6231	6 AR302610	AR302610 Sequence
10	10451.5	96.7	6510	6 AR302607	AR302607 Sequence
11	10435	96.5	6812	6 AB002342	AB002342 Human mRN
12	10375	96.0	6969	6 AR302604	AR302604 Sequence
13	10344	95.7	6147	6 AR302618	AR302618 Sequence
14	10287.5	95.1	6426	6 AR302615	AR302615 Sequence
15	10211	94.4	6885	6 AR302612	AR302612 Sequence
16	9929	91.8	6000	6 AR302595	AR302595 Sequence
17	9872.5	91.3	6279	6 AR302589	AR302589 Sequence
18	9796	90.3	6718	6 AR302603	AR302603 Sequence
19	9765	90.6	5916	6 AR302603	AR302603 Sequence
20	9729.5	90.0	7328	6 AX056418	AX056418 Sequence
21	9708.5	89.8	6195	6 AR302600	AR302600 Sequence
22	9632	89.1	5820	6 AR302597	AR302597 Sequence
23	9627	89.0	5820	6 AR302611	AR302611 Sequence
24	9570.5	88.5	6099	6 AR302608	AR302608 Sequence
25	9494	87.8	6558	6 AR302605	AR302605 Sequence
26	9463	87.5	5736	6 AR302619	AR302619 Sequence
27	9406.5	87.0	6015	6 AR302616	AR302616 Sequence
28	9330	86.3	6474	6 AR302613	AR302613 Sequence
29	9204	85.1	7342	10 AF227741	AF227741 Rattus no
30	9041.5	83.6	9313	10 AR309076	AR309076 Mus muscu
31	4053	37.5	2949	6 AR302590	AR302590 Sequence
32	4043	37.4	2490	6 AR302593	AR302593 Sequence
33	3889	36.0	2865	6 AR302598	AR302598 Sequence
34	3879	35.9	2406	6 AR302601	AR302601 Sequence
35	3783.5	35.0	5061	10 BC056761	BC056761 Mus muscu
36	3751	34.7	2769	6 AR302606	AR302606 Sequence
37	3741	34.6	2310	6 AR302609	AR302609 Sequence
38	3587	33.2	2685	6 AR302614	AR302614 Sequence
39	3577	33.1	2226	6 AR302617	AR302617 Sequence
40	3408	31.5	2653	9 AC004803	AC004803 Homo sapi
41	3107	28.7	154369	9 AC004803	AC004803 Homo sapi
42	2953	27.3	2310	6 AX224728	AX224728 Mus muscu
43	2844	26.3	2640	10 BC055788	BC055788 Homo sapi
44	2638.5	24.4	6891	9 HSA242724	AJ242724 Homo sapi
45	2621	24.2	7792	9 AB051547	AB051547 Homo sapi

RESULT 1

ALIGNMENTS

AR302594 AR302594 6411 bp DNA linear PAT 12-JUN-2003
 LOCUS AR302594
 DEFINITION Sequence 13 from patent US 6541252.
 ACCESSION AR302594
 VERSION AR302594.1 GI:31690892
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE Unclassified.
 1 (bases 1 to 6411)
 AUTHORS Walke, D.W., Hilburn, E., Donoho, G. and Turner, C.A. Jr.
 TITLE Human kinases and polynucleotides encoding the same
 JOURNAL Patent: US 6541252-A 13 01-APR-2003;
 FEATURES
 source 1. 6411
 /organism="unknown"
 /mol_type="genomic DNA"

Alignment Scores:
 Pred. No.: 5.52e-258 Length: 6411
 Score: 10810.00 Matches: 2136
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 99.98% Indels: 0
 DB: Gaps: 0

US-10-010-720-14 (1-2136) x AR302594 (1-6411)

QY 1 MetSerGlyGlyAlaAlaGluValGlnSerSerThrProGlySerLeuPheLeuSerPro 20
 Db 1 ATGTCTGGCGCGCGCGAGAGACAGACAGACACTCCGGTTCCTGTTCTCTCCCGC 60
 QY 21 ProAlaProAlaProLysAsnGlySerSerSerSerSerValGlyGlyLeuGly 40
 Db 61 CCGGCTCTGCGCGCGCGAGAGATGGCTCCAGCTCCGATTCCTCCGCGGAGAACTGGGA 120
 QY 41 ALAAlaAlaAlaAlaAlaValThrGlyArgThrGluGluValArgAlaGlyArgHisThr 60
 Db 121 GCCCGCGCGCGCGCGGTGTGACCGGACGACCGAGAGAGACAGCGCGCGCGCACACT 180
 QY 61 MetAspLysAspSerSerArgGlyAlaAlaAlaThrThrThrThrGluHisArgPhePhe 80
 Db 181 ATGACAGAGAGACAGCGCTGGGCGCGCGCGACCACTACCACTGAGCAGCGCTTCTTC 240
 QY 81 ArgArgSerValIleCysAspSerAsnAlaThrAlaLeuGluLeuProGlyLeuProLeu 100
 Db 241 CGCGGAGCGCTCACTGTGCACTCCAAATGCCACTGCACCTGAGCTTCCCGGCTTCTCTT 300
 QY 101 SerLeuProGlnProSerIleProAlaAlaValProGlnSerAlaProProGluProHis 120
 Db 301 TCCCTGGCGCGCGCGAGCATCCCGCGCTGTCCCGAGAGTGTCCACCGAGCGCGCCAC 360
 QY 121 ArgGluGluThrValThrAlaThrAlaThrSerGlnValAlaGlnGlnProProAlaAla 140
 Db 361 CGGAGAGAGACCGTGCACCGCCACTTCCAGGTAGCCGACGACCTCCAGCCCT 420
 QY 141 ALAAlaProGlyGluGlnAlaValAlaGlyProAlaProSerThrValProSerSerThr 160
 Db 421 GCCCGCGCGCGCGAGACAGCGCTGGCGGCGCTGCCCTCGACTGTCCCGAGAGTACC 480
 QY 161 SerLysAspArgProValSerGlnProSerLeuValGlySerLysGluGluProProPro 180
 Db 481 AGCAAGAGCGCGCGAGTGTCCAGGCTAGCTTGTGGGAGCAAGAGAGACCGCGCGC 540
 QY 181 AlaArgSerGlySerGlyGlyGlySerAlaLysGluProGlnGluGluArgSerGlnGln 200
 Db 541 GCGAGAAATGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 600
 QY 201 GlnAspAspIleGluGluLeuGluThrLysAlaValGlyMetSerAsnAspGlyArgPhe 220
 Db 601 CAGATGATATCGAAGAGCTGAGACCAAGCGCGCGGAGAAATGTCTAAAGATGCGCTTT 660

QY 221 LeuLysPheAspIleGluIleGlyArgGlySerPheLysThrValTyrLysGlyLeuAsp 240
 Db 661 CTCAGATTGACATCGAAGATCGCGAGAGGCTCTTTAAGCGGTCTACAAAGGTCGGAC 720
 QY 241 ThrGluThrThrValGluValAlaTyrCysGluLeuGlnAspArgLysLeuThrLysSer 260
 Db 721 ACTGAACACCGGTGAGAGCTGCTGTGTGACATCGAGATCGAAATTAACAAAGTCT 780
 QY 261 GluArgGlnArgPheLysGluGluAlaGluMetLeuLysGlyLeuGlnHisProAsnIle 280
 Db 781 GAGAGGACAGAGATTAAAGAGAGAGTGAATGTAAAGGCTCTTACGATCCCAATATT 840
 QY 281 ValArgPheTyrAspSerTyrProGlnSerThrValLysGlyLysLysCysIleValLeuVal 300
 Db 841 GTTAGATTATATATCTCTGGGAATCCAGATTAAGAGAGAGAGAGAGAGAGAGAGAG 900
 QY 301 ThrGluLeuMetThrSerGlyThrLeuLysThrTyrLeuLysArgPheLysValMetLys 320
 Db 901 ACTGAACCTATGACGTCTGGAGACCTTAACCACTTAACCTATCTGAAGGTTTAAAGTATGAG 960
 QY 321 IleLysValLeuArgSerTyrCysArgGlnIleLeuLysGlyLeuGlnPheLeuHisThr 340
 Db 961 ATCAAGTTCTTAAGAGAGCTGTCGCGACATCTTAAGGCTCTTCACTTCTTCACTACT 1020
 QY 341 ArgThrProProIleIleHisArgAspLeuLysCysAspAsnIlePheIleThrGlyPro 360
 Db 1021 CGAATCTCACTTATCATCTTACCGCGATCTTAATGTACAAACATCTTTACCGCGCTCT 1080
 QY 361 ThrGlySerValLysIleGlyAspLeuGlyLeuAlaThrLeuLysArgAlaSerPheAla 380
 Db 1081 ACTGGCTCACTCAAGATTGAGACCTCGCTGTGCAACCTGAAAGCGGCTCTCTTTGTC 1140
 QY 381 LysSerValIleGlyThrProGluPheMetAlaProGluMetTyrGluGluLysTyrAsp 400
 Db 1141 AAGAGTGTATAGTATCCCAAGATTCATGCGCGCGAGATGATGAGAGAGAGATATGAT 1200
 QY 401 GluSerValAspValTyrAlaPheGlyMetCysMetLeuGluMetAlaThrSerGluTyr 420
 Db 1201 GAATCCCTTACGTTATATGCTTTTGGAGTGTGATGCTTGAATGCTTACATCTGATATAT 1260
 QY 421 ProTyrSerGluCysGlnAsnAlaAlaGlnIleTyrArgArgValThrSerGlyValLys 440
 Db 1261 CTTACTCGGAGTCCCAAAATGCTGCACAGATCATCCGTGCGGACAGTGGGTTAAG 1320
 QY 441 ProAlaSerPheAspLysValAlaIleProGluValLysGluIleIleGluGlyCysIle 460
 Db 1321 CCAAGCGATTTTGAAGAAAGTGAATTCGAAATGAGAAATTAATTAAGATGATCA 1380
 QY 461 ArgGlnAsnLysAspGluArgTyrSerIleLysAspLeuLeuAsnHisAlaPhePheGln 480
 Db 1381 CGACAAACAAAGATGAAGATATTCATCAAGACCTTTGAACCATGCTTCTTCCAA 1440
 QY 481 GluGluThrGlyValArgValGluLeuAlaGluGluAspAspGlyGluLysIleAlaIle 500
 Db 1441 GAGAGAAACAGAGATGCGGTGAATTAAGCAAGAAATGATGAGAGAGAGAGAGAGAG 1500
 QY 501 LysLeuThrLeuAlaGluAspIleLysLysLeuLysGlyLysTyrLysAspAsnGlu 520
 Db 1501 AAATTAATGCTACATATTAAGATATTAAGAAATTAAGAGAAATTAAGAAATTAAGAA 1560
 QY 521 AlaIleGluPheSerPheAspLeuGluArgAspValProGluAspValAlaGlnGluMet 540
 Db 1561 GCTATTTAGATTCTTTTATTGATTAGAGAGAGATGCCAGAGAGATGTGCAAGAAATG 1620
 QY 541 ValGluSerGlyTyrValCysGluGlyAspHisLysThrMetAlaLysIleLysAsp 560
 Db 1621 GTAGAGTCTGGGTATGCTGTGAAGGTGATCAAGACAGAGGTAAAGCTATTAAGAGAC 1680
 QY 561 ArgValSerLeuIleLysArgLysArgGluGlnArgGluMetValArgGluGluGlu 580
 Db 1681 AGAGTATCATTAATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1740
 QY 581 LysLysLysGlnGluLysSerSerLeuLysGlnGlnValGluGlnSerSerAlaSerGln 600


```
Db 1741 AAAAAAAAAAGCAGAGAGAGAGCTCTCAACAGCAGGTGGAACATCCAGTCTTCCAG 1800
Qy 601 ThrGlyIleLysGlnLeuProSerAlaSerThrGlyIleProThrAlaSerThrThrSer 620
Db 1801 AACAGAAATCAGACAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCA 1860
Qy 621 AlaSerValSerThrGlnValGlnProGluGluProGluAlaAspGlnHisGlnGlnLeu 640
Db 1861 GCTTACAGTTTCTTCAACAAAGTAAACCTTCAAGAACTTGAGGAGCATCAACATCAACACTA 1920
Qy 641 GlnIyrGlnGlnProSerIleSerValLeuSerAspGlyThrValAspSerGlyGlnGly 660
Db 1921 CAGTACCAAGCAACCCAGATATCTGTGTATCTGATGGAACGTTGAGACAGTGTCAAGGA 1980
Qy 661 SerSerValPheThrGlnSerArgValSerSerGlnGlnThrValSerTyrGlySerGln 680
Db 1981 TCCTCTGCTCTCAACAGATCTGAGTGAAGCAGCCCAACAGACAGTTTCAATGCTTCCAA 2040
Qy 681 HisGlnGlnAlaHisSerThrGlyThrValProGlyHisIleProSerThrValGlnAla 700
Db 2041 CATGAACAGGACATTTTCAAGGACAGCTCCAGGAGCATATACCTTCTACTGTCACAGCA 2100
Qy 701 GlnSerGlnProHisGlyValTyrProProSerSerValAlaGlnGlyGlnSerGlnGly 720
Db 2101 CAGTCTACGCCCATGGGTATATCCACCTCAAGTGTGAGCAGAGGAGCAGAGCCAGGCT 2160
Qy 721 GlnProSerSerSerSerLeuThrGlyValSerSerSerGlnProIleGlnHisProGln 740
Db 2161 CAGCATCTCTCAAGTACCTTAAACAGGGGTTTCATCTTCCAAACCATACATCCCTCAG 2220
Qy 741 GlnGlnGlnGlnGlyIleGlnGlnThrAlaProProGlnGlnThrValGlnIyrSerLeuSer 760
Db 2221 CAGAGACAGGAAATACAGCAGACAGCCCTCTCTCAACAGACAGTACCTCAAGCTCCA 2280
Qy 761 GlnThrSerThrSerSerGlnAlaThrThrAlaGlnProValSerGlnProGlnAlaPro 780
Db 2281 CAGACATCAACCTCCAGTACAGGACCACTACCTGACAGCCAGTACAGCTCAAGCTCCA 2340
Qy 781 GlnValLeuProGlnValSerAlaGlyLysGlnSerThrGlnGlyValSerGlnValAla 800
Db 2341 CAAGTCTTGCCTCAAGTATCAGCTGGAACACAGAGTCTCAGGAGTCTCAGGTTGCT 2400
Qy 801 ProAlaGlnProValAlaValAlaAlaGlnProGlnAlaThrGlnProThrThrLeuAlaSer 820
Db 2401 CCTGCAGAGCCAGTGTGAGTAGCAGACACCCCAAGCTTCCAGCCAGCACCTTGGCTTCC 2460
Qy 821 SerValAspSerAlaHisSerAspValAlaSerGlyMetSerAspGlyAsnGlnValSer 840
Db 2461 TCTGTACACGTGCACATTCAGATGTTGCTTCAAGTATGAGTAGCAATGGAACGCTC 2520
Qy 841 ProSerSerSerGlyArgHisGlnGlyArgThrThrLysArgHisTyrArgLysSerVal 860
Db 2521 CCATCTTCCAGTGAAGGACATGAAGAGAACACACAAAGGATTAACGAAATCTCTGA 2580
Qy 861 ArgSerArgSerArgHisGlyLysThrSerArgProLysLeuArgIleLeuAsnValSer 880
Db 2581 AGGAGTGTGCTCTGAGATGAAAAAACTTCAAGCCCAAAATTAAAGAAATTTGATGTTTCA 2640
Qy 881 AsnLysGlyAspArgValAlaGlnCysGlnLeuGlnThrHisAsnArgLysMetValThr 900
Db 2641 AATTAAGAGACCGAGTATGATGATCTCAATTAAGAACTCAATTAATGAGAAAAAGGTTTAA 2700
Qy 901 PheLysPheAspLeuAspGlyAspAsnProGlnGlnIleAlaThrIleMetValAsnAsn 920
Db 2701 TTCAATTTGACCTAGATGCTGACACCCGAGGAGATAGCAACATTTATGGAACAT 2760
Qy 921 AspheIleLeuAlaIleGlnArgGlnLeuSerPheValAspGlnValArgGlnIleIleGln 940
Db 2761 GACTTATTTCTGCAATTAAGAGAGAGTGTGTTGTGATCAAGTGCAGAAATTTATTTGAA 2820
Qy 941 LysAlaAspGlnMetLeuSerGlnAspValSerValGlnProGlnGlyLysGlnGlyLeu 960
Db 2821 AAAGCTGATGAATGCTCAGTGAAGATGTCAGTGTGAAACAGAGGATGATCAGGAGATTG 2880
Qy 961 GluSerLeuGlnGlyLysAspAspTyrGlyPheSerGlySerGlnLysLeuGlnGlyGln 980
Db 2881 GAGGCTTACAGAGAAAGAAAGATGACTATGAGCTTTTTCAGAGTTTCCAGAAATTTGAGAGAGAG 2940
Qy 981 PheLysGlnProIleProAlaSerSerMetProGlnGlnIleGlyIleProThrSerSer 1000
Db 2941 TTCAACAGACCAATTTCTCGCTCTTCCATGCGCACAGCAAAATAGGCATTTCCAGCTT 3000
Qy 1001 LeuThrGlnValAlaHisSerAlaGlyArgArgPheIleValSerProValProGluSer 1020
Db 3001 TTAATCAAGTTGTTCATTTGCGGAAAGGGGTTTATAGAGATCTGTCTCCAGAAAGC 3060
Qy 1021 ArgLeuArgGlnSerLysValPheProSerGlnIleThrAspThrValAlaIleSerThr 1040
Db 3061 CGATTACAGAGATCAAAAGTTTCCAGTAAATTAACAGATACAGTTGCTGCTCTTCA 3120
Qy 1041 AlaGlnSerProGlyMetAsnLeuSerHisSerAlaSerSerLeuSerLeuGlnAla 1060
Db 3121 GCTCAGAGCCCTGGAAATGAATCTTCTCATCTGATCATCTTCTTACTTCAACAGAGCC 3180
Qy 1061 PheSerGlnLeuArgArgAlaGlnMetThrGlnGlyProAsnThrAlaProProAsnPhe 1080
Db 3181 TTTTCTGAACCTTAGACGTGCCAAATGACAGAAAGACCAATACAGACCTCCAAACTTT 3240
Qy 1081 SerHisThrGlyProThrPheProValAlaProProPheLeuSerSerIleAlaGlyVal 1100
Db 3241 AGTCATACAGAGACCAATTTCCAGTACGTACCTCTCTTCTTCAAGTACATTTGCTGAGTC 3300
Qy 1101 ProThrThrAlaAlaAlaThrAlaProValProAlaThrSerSerProProAsnAspIle 1120
Db 3301 CCAACCAACAGACAGGACCAAGCACAGCAGCTCCCTGCAACAGACCCCTCTCAATGACATT 3360
Qy 1121 SerThrSerValIleGlnSerGlnValThrValProThrGlnGlnGlyIleAlaGlyVal 1140
Db 3361 TCCACATCACTAATTAATTCAGTGTGAGTTTACAGTGCCCATGAAGGGGATTTGCTGAGTT 3420
Qy 1141 AlaThrSerThrGlyValValThrSerGlyGlyLeuProIleProProValSerGluSer 1160
Db 3421 GCCACCAAGCACAGGTGTGTAATTCAGGTGCTCCCAATACACCTGTGTCTGAATCA 3480
Qy 1161 ProValLeuSerSerValAlaSerSerIleThrIleProAlaValAlaSerIleSerThr 1180
Db 3481 CCAAGTACTTCCAGCTGATGTTCAAGTATACAAATCCGCAATGTGCTCAATATCTACT 3540
Qy 1181 ThrSerProSerLeuGlnValProThrSerThrSerGlnIleValAlaSerSerThrAla 1200
Db 3541 ACATCCCCGTCACCTTCAAGTCCCAATCCCAATCCATCGAATCGTTGTCTTACTATACGA 3600
Qy 1201 LeuTyrProSerValThrValSerAlaThrSerAlaSerAlaGlySerThrAlaThr 1220
Db 3601 CGETATCTTCAAGTAAACATTTTCAGCACTTCAAGCTCTGCAAGGGGAGTACTGCTTACC 3660
Qy 1221 ProGlyProLysProProAlaValAlaSerGlnGlnAlaAlaGlySerThrThrValGly 1240
Db 3661 CCAAGTCTCAAGCTCCAGCTGATGATCTCAGACAGGAGGAGGAGCACTACTGTGGGA 3720
Qy 1241 AlaThrLeuThrSerValSerThrThrThrSerPheProSerThrAlaSerGlnLeuSer 1260
Db 3721 GCCACATTTAAACATGATTTCTACACCACTTCAATCCCAAGCAGCTTCAAGCTGTGC 3780
Qy 1261 IleGlnLeuSerSerSerThrSerThrProThrLeuAlaGlnThrValValAlaSerAla 1280
Db 3781 ATTCAGCTTAGAGCAGTACTTCTACTCTTAACTTGAAGAAACCGTGTAGTTAGCGCA 3840
Qy 1281 HisSerLeuAspLysThrSerHisSerSerThrThrGlyLeuAlaPheSerLeuSerAla 1300
Db 3841 CACTCACTAGATTAAGACATCTATACAGTACATCAATGCAATTTGCTTCTCTCTGGA 3900
Qy 1301 ProSerSerSerSerSerProGlyAlaGlyValSerSerTyrIleSerGlnProGlyGly 1320
Db 3901 CCATCTTCTCTCTCTCTCTCTGAGCAGAGAGTGTCAATTAATTTCTCAGCTGTGTGG 3960
```

QY 1321 LeuHisProLeuValIleProSerValIleIleSerThrProIleLeuProGlnIleAla 1340
| | | | |
Db 3961 CTGCATCCCTTGGTCAATTCATCAGTGAAGTCTTCACTCTTATCTTCCCAAGACGA 4020
| | | | |
QY 1341 GlyProThrSerThrProLeuLeuProGlnValProSerIleProProLeuValGlnPro 1360
| | | | |
Db 4021 GGAACCTACTCTACACTTATTAACCCAAAGTACTAGTATCCCACTTGGTGAACGCT 4080
| | | | |
QY 1361 ValAlaAsnValProAlaValAlaGlnGlnThrLeuIleHisSerGlnProGlnProAlaLeu 1380
| | | | |
Db 4081 GTTGCCAAATGGCTGCTGTACAGACAGACATAATTATAGTCAGCTCAACAGCTTGG 4140
| | | | |
QY 1381 LeuProAsnGlnProHisThrHisCysProGluValAspSerAspThrGlnProValAla 1400
| | | | |
Db 4141 CTTCGCCAACAGCCCTTACTCATTTGCTGAAGTATCTGTATACCAACCCAAAGCT 4200
| | | | |
QY 1401 ProGlyTyleAspAspIleLeuThrLeuGlnGlnLeuArgSerLeuPheSerGlnHis 1420
| | | | |
Db 4201 CTGGAATTATGACATTAAGACTTGAAGAAAGCTGGGCTCTGTTCAAGTGAAC 4260
| | | | |
QY 1421 SerSerSerGlyAlaGlnHisAlaSerValSerLeuGlnThrSerLeuValIleGlnSer 1440
| | | | |
Db 4261 AGCTCATCTGGAGTCAAGATGCTCTGTCTCACTGAGAGCTCACTAGTCATAGAGAGC 4320
| | | | |
QY 1441 ThrValThrProGlyIleProThrThrAlaValAlaProSerIleLeuLeuThrSerThr 1460
| | | | |
Db 4321 ACTGTCAACACAGGCACTCCCAACTACTGCTGTCCACCAAGCAAACTCTGACTTCAAC 4380
| | | | |
QY 1461 ThrSerThrCysLeuProProThrAsnLeuProLeuGlyThrValAlaLeuProValThr 1480
| | | | |
Db 4381 ACAAGTACTTGCTTACCAACCAAAATTTACCACTAGGAACAGTTGCTTCCAGTTACA 4440
| | | | |
QY 1481 ProValValThrProGlyGlnValSerThrProValSerThrThrThrSerGlyValIle 1500
| | | | |
Db 4441 CCAATGGTCAACCTGGGCAAGTTTCTTACCCAGTCAAGCACTACTACATCAGAGTGA 4500
| | | | |
QY 1501 ProGlyThrAlaProSerLeuProProLeuThrIleValAlaProValLeuProValGlyThr 1520
| | | | |
Db 4501 CCTGAATGCTCTCCCTCCAAAGCACTCTTAAGGCTCGAGTCCAGTGGTACT 4560
| | | | |
QY 1521 GlnLeuProAlaGlyThrLeuProSerGlnGlnLeuProProPheProGlyProSerLeu 1540
| | | | |
Db 4561 GAACCTTCAGAGGTACTCTACCCAGCAGACGCTGCCACTTTTCCAGAACCTTCTCTA 4620
| | | | |
QY 1541 ThrGlnSerGlnGlnProLeuGlnAspLeuAspAlaGlnLeuArgGlyThrLeuSerPro 1560
| | | | |
Db 4621 ACCCAAGTCCAGCAACCTCTAAGAGATCTGTATGCTCAATTGAGAAACAATTAGTCA 4680
| | | | |
QY 1561 Glnu***IleThrValThrSerAlaValGlyProValSerMetAlaAlaProThrAlaIle 1580
| | | | |
Db 4681 GAGATKACACAGTACTTGGCGGTGTCTGTGTCCATGGCGCTCCAAACAGCAATC 4740
| | | | |
QY 1581 ThrGlnIleGlyThrGlnProGlnIleGlyValSerGlnValIleGlyValProValLeu 1600
| | | | |
Db 4741 ACAGAAAGGAAACACAGCTTCAGAAAGGTGTTCTCAAGTCAAGAAAGGCCCTCTCTA 4800
| | | | |
QY 1601 AlaThrSerSerGlyAlaGlyValAlaPheLeuSerGlyArgPheGlnValSerValAlaAla 1620
| | | | |
Db 4801 GCAACTAGTTCAGAGGCTGTGTATTAAGATGGAGCATTTCAAGTTCTGTTCCAGGA 4860
| | | | |
QY 1621 AspGlyAlaGlnLeuGlyValIleValAsnLeuSerGlnAspAlaIleValSerValHisPheGln 1640
| | | | |
Db 4861 GACGGTGGCCAGAAAGAGGGTAAAAATAGTCAAGAGATGCAAGAGCTGTGTAATTTGGA 4920
| | | | |
QY 1641 SerSerThrSerGlnSerSerValLeuSerSerSerSerProGlnSerThrThrValIle 1660
| | | | |
Db 4921 TCCAGACCTCAAGAGCTCTCAAGTCTATCAAGTATAGTCCAGAGATACCTTGGTGA 4980
| | | | |
QY 1661 ProGlnProAsnGlyIleThrIleProGlyIleSerSerAspValProGlnSerAlaHis 1680
| | | | |
Db 4981 CCAAGGCCAAAGGCAATACATCCTGCTATCTTCAAGATGTGCCAGAGATGCCCCAC 5040
| | | | |

QY 1681 LysThrThrAlaSerGlnAlaIleSerSerAspThrGlnProThrIleValGlyArgPhe 1700
| | | | |
Db 5041 AAACATCTGCTCCAGAGGCAAGTACAGACATGGGCAAGCTTACCAAGGTTGACCTTTT 5100
| | | | |
QY 1701 GlnValThrThrThrAlaAsnLeuValGlyArgPheSerValSerIleThrGlnAspIle 1720
| | | | |
Db 5101 CAGGTACACACTCAACCAAAAGAGGGTGTCTTCTGTATCAAAAATGAGAGCAAG 5160
| | | | |
QY 1721 IleThrAspThrIleValGlyGlyProValAlaSerProPheMetAspLeuGlnIle 1740
| | | | |
Db 5161 ATCACTGACCAAAAGAAAGAGGACAGTGGCACTCTCTCTTTATGATTTGGAAACA 5220
| | | | |
QY 1741 AlaValLeuProAlaValIleProIleGlyGlyLeuAspGlnLeuSerGlnProSerHis 1760
| | | | |
Db 5221 GCTGTTCTTCTGCTGTGATACCAAAAGAAAGAGCTTCACTTATGATTTGGAAACA 5280
| | | | |
QY 1761 LeuAsnGlyProSerSerAspProGlnAlaAlaPheLeuSerArgAspValAspAspGly 1780
| | | | |
Db 5281 CTAATGGGCGCTCTCTGACCCGAGGCGCTTTTAAAGTGGAGTGTGATGATGCT 5340
| | | | |
QY 1781 SerGlySerProHisSerProHisGlnLeuSerSerIleSerLeuProSerGlnAsnLeu 1800
| | | | |
Db 5341 TCCGATGTCACACTGCCCCATCACTAGCTCAAAAGAGCTTCTTACCCAGAACTA 5400
| | | | |
QY 1801 SerGlnSerLeuSerAsnSerPheAsnSerSerIleMetSerSerAspAsnGlnSerAsp 1820
| | | | |
Db 5401 AGTCAAAGCTTGAATTCATTTAATCTCTTACCTGATGAGTGCAGCAATGAGTCAAT 5460
| | | | |
QY 1821 IleGlnAspGlnAspLeuValIleIleProProAlaAlaProLeuSerGlyArgArgArg 1840
| | | | |
Db 5461 ATCGAATGATGAAGCTTAAAGTTAGAGCTCGAGATGACATGATTAACATCTCAAGAG 5520
| | | | |
QY 1841 IleGlnAspLeuGlnSerArgGlnIleHisGlnIleGlnSerLeuThrIleValGly 1860
| | | | |
Db 5521 ATTGAGACCTGACAGATCCGCCACAGACATGAATTTGAATCTTGTATCCAAACTGGGC 5580
| | | | |
QY 1861 LysValProProAlaValIleIleProProAlaAlaProLeuSerGlyArgArgArg 1880
| | | | |
Db 5581 AAGGTGCCCTGCTGTATATATCCCCAGCTCTCTCTTACAGGAGAAAGCAGAGA 5640
| | | | |
QY 1881 ProThrIleSerIleGlySerIleSerSerSerSerSerSerSerLeuGlyAsnIleSerPro 1900
| | | | |
Db 5641 CCCACTAAAGCAAAAGCAGCAAAATCTAGTGAAGCAAGTTCTTGGGAAATAAAGCCCC 5700
| | | | |
QY 1901 GlnLeuSerGlyAsnLeuSerGlyGlnSerAlaIleSerValIleHisProGlnGlnThr 1920
| | | | |
Db 5701 CAGCTTTCAGATTAACCTGTCTGTCTCAGAGTCACTTCAAGTCTTGCACCCCAAGACAC 5760
| | | | |
QY 1921 LeuHisProProGlyAsnIleProGlnSerGlyGlnAsnGlnLeuLeuGlnProLeuIle 1940
| | | | |
Db 5761 CTCACACCTCTGGCAACATCCAGAGTCCGGGCAAGATCAGCTGTATACAGCCCTTTAAG 5820
| | | | |
QY 1941 ProSerProSerSerAspAsnLeuThrSerAlaPheThrSerAspGlyAlaIleSerVal 1960
| | | | |
Db 5821 CCATCTCCCTCCAGAGCAACCTCTATTCACTTCACTTCACTGATGATGTCATTTCACTA 5880
| | | | |
QY 1961 ProSerLeuSerAlaProGlyGlnGlyThrSerSerThrAsnThrValGlyAlaThrVal 1980
| | | | |
Db 5881 CCAGACCTTTCTGCTCCAGGTCAGAGAACAGACAGACAAACACTGTTGGGCAACAGTG 5940
| | | | |
QY 1981 AsnSerGlnAlaAlaGlnAlaGlnProProAlaMetThrSerSerArgIleGlyThrPhe 2000
| | | | |
Db 5941 AACAGCCAGCCGCCCAAGCTCAGCTCTGCTCATGAGCTCCAGAGAGAAAGGCACTTC 6000
| | | | |
QY 2001 ThrAspAspLeuHisIleValLeuValAspAsnThrAlaArgAspAlaMetAsnLeuSerGly 2020
| | | | |
Db 6001 ACAGATGACTTGCAACAGTGTGATGACCAATGGGCCCCAGATGACATGAATTTCTCAAGC 6060
| | | | |
QY 2021 ArgArgGlySerIleGlyHisMetAsnThrGlnGlyProGlyMetAlaArgIlePheSer 2040
| | | | |
Db 6061 AGGAGAGGAGCAAGAGGCACTGATTTATGAGGCGCTCGGAATGGCAAGAAATGCTCT 6120
| | | | |
QY 2041 AlaProGlyGlnLeuCysIleSerMetThrSerAsnLeuGlyIleSerAlaProIleSer 2060
| | | | |

QY 441 ProAlaSerPheAspIysValAlaIleProGluValIleGluIleIleGluIysCysIle 460
DB 1739 CCGAGCAAGTTTTCACAAAGATGCAATTCCTGAGTAAAGAAATTTATGAAGATGCATA 1798
QY 461 ArgGlnAsnLysAspGluArgIysSerIleYAspLeuAsnHisAlaPhePheGln 480
DB 1799 CGACAAAACAAAGTGAAGATATTCATCAAGAACCTTTTGAACCAATGCTTCCCA 1858
QY 481 GluGluThrIleValArgValGluLeuAlaGluGluAspAspGluIleAlaIle 500
DB 1859 GAGGAAACAGAGATACGGGTAGAAATTTAGCAAGAAAGATGATGAGAAAAAATAGCCATA 1918
QY 501 LysLeuThrLeuArgIleGluAspIleLysLysLeuLysGlyLysThrLysAspAsnGlu 520
DB 1919 AAATTATGCTAGCTATTAAGAAATTTAAAGGAAATTAACAAAGTAAATGAA 1978
QY 521 AlaIleGluPheSerPheAspLeuArgAspValProGluAspValAlaGluIleMet 540
DB 1979 GCTATTAGTTTCTTTTGAATTTAGAGAGATGCTCCAGAAAGATGTTCACAAAGAAATG 2038
QY 541 ValGluSerGlyIysValCysGluGlyAspHisLysThrMetAlaLysAlaIleLysAsp 560
DB 2039 GTAGAGTCTGGGTATGCTGTGAGAGTGATCACAAGACCATGGCTAAAGCTATCAAGAC 2098
QY 561 ArgValSerLeuIleLysArgLysArgGluGluArgGluMetValArgGluGluGlu 580
DB 2099 AGAGTATCATTAATTAAGAGAAACGAGAGCGGCGATTGTGACGGAGAGCAAGAA 2158
QY 581 LysLysLysGluGluGluSerSerLeuLysGluIleValGluGluSerSerAlaSerGln 600
DB 2159 AAAAAAAGCAGAGAGAGAGAGAGCTCTAAACAGCAGTAAACATCAAGTCTCCAG 2218
QY 601 ThrGlyIleLysGluLeuProSerAlaSerThrGlyIleProThrAlaSerThrThSer 620
DB 2219 ACAGAAATCAAGAGAGCTCTCTGCTAGACCGGCAATCACTACTGCTTCAACACTTCA 2278
QY 621 AlaserValSerThrGluValGluProGluGluProGluAlaAspGlnHisGluGluLeu 640
DB 2279 GCTTCAGTTTTCACAAAGTAAACCTGAAGACCTGAGGCAATCAATCAACACTA 2338
QY 641 GlnIysGluGlnProSerIleSerValLeuSerAspGlyThrValAspSerGlyGlnGly 660
DB 2339 CAGTACAGAGCAACCAAGTATCTGTATCTGATGGGACGGTTGACAGTGTCAAGGA 2398
QY 661 SerSerValPheThrGluSerArgValSerSerGluGlnThrValSerThrGlySerGln 680
DB 2399 TCCTGTGCTTCAAGAAATCTGAGTGAAGCCAAACAGACAGTTTCAATATGTTCCCA 2458
QY 681 HisGluGlnAlaHisSerThrGlyThrValProGluHisIleProSerThrValGlnAla 700
DB 2459 CATGACAGGCAATTTACAGGCAAGCTCCAGGCAATATACCTTCTACTCTCCAAAGCA 2518
QY 701 GlnSerGlnProHisGlyValIysProProSerSerValAlaGlnGlyGlnSerGlnGly 720
DB 2519 CAGTTCACGCCCAAGGGGTATATCCACCTCAAGTGTGGCAAGGGGAGAGCCAGGT 2578
QY 721 GlnProSerSerSerSerLeuThrGlyValSerSerSerGlnProIleGlnHisProGln 740
DB 2579 CAGCCATCTCTCAAGTATGCTTAAACAGGGGTTTCATCTTCCCAACCAATACATCCT 2635
QY 741 GlnGlnGlnGlyIleGlnGlnThrAlaProProGlnGlnThrValGlnIysSerLeuSer 760
DB 2636 CAGCAGCAGGAAATACAGACAGACGCCCTCTCCACAGACAGTGCAGATTCACCTTCA 2695
QY 761 GlnThrSerThrSerSerGluAlaThrThrAlaGlnProValSerGlnProGlnAlaPro 780
DB 2696 CAGACATCAACCTCCAGTGAAGCCACTACTGCACACCAAGTAGTCAACTCAAGCTCCA 2755
QY 781 GlnValLeuProGlnValSerAlaGlyLysGlnSerThrGlnGlyValSerGlnValAla 800
DB 2756 CAAGTCTTGCTCAAGTATCAAGTGAAGAAACAGAGTACTCAGGAGATCTCTCAGGTGCT 2815
QY 801 ProAlaGluProValAlaValAlaGlnProGlnAlaThrGlnProThrThrLeuAlaSer 820

DB 2816 CCGCAGAGCAGCTTTCAGTACACAGCCCAAGCTAACCCAGCCGACCTTGGCTTCC 2875
QY 821 SerValAspSerAlaHisSerSerValAlaSerGlyMetSerAspGlyLysGluMetVal 840
DB 2876 TCTGTAGACAGTGCACATTCAGTGTGCTTCAAGTATGATGATGATGATGATGATGATG 2935
QY 841 ProSerSerSerGlyArgHisGluGlyArgThrThrLysArgHisIysArgLysSerVal 860
DB 2936 CCATCTTCAGTGAAGGATGAGAGAAATACAAACGCGATTACCGAAATATCTGTA 2995
QY 861 ArgSerArgSerArgHisGluLysThrSerArgProLysLeuArgIleLeuAsnValSer 880
DB 2996 AGAGTCTGCTCGACATGAAAAAACTTCAAGCCCAAAATTAAGAAATTTGAATGTTC 3055
QY 881 AsnLysGlyAspArgValValGluCysGlnLeuGluIleThrHisAsnArgLysMetValThr 900
DB 3056 AATAAAGAGACCGAGTATGAAATGTCAATTTAGAACTATTAAGAAAAATGTTTACA 3115
QY 901 PheLysPheAspLeuAspGlyAspAsnProGluGluIleAlaThrIleMetValAsnAsn 920
DB 3116 TTCAATTTGACCTAGATGTTGACAAACCCAGAGATAGCAACAAATTATGTTGAAACAT 3175
QY 921 AspPheIleLeuAlaIleGluArgGluSerPheValAspGlnValArgGluIleIleGlu 940
DB 3176 GACTTATCTTACGAAATGAGAGAGAGTCTTGTGATCAAGTGCAGAAATTTATTGAA 3235
QY 941 LysAlaAspGluMetLeuSerGluAspValSerValGluProGluGlyAspGlnGlyLeu 960
DB 3236 AAGCTATGAAATGCTCAAGTGAAGATGTCAAGTGTGAACCAAGGGTGTACGGAGTTG 3295
QY 961 GluSerLeuGlnGlyLysAspAspIysGlyPheSerGlySerGlnLysLeuGluGlu 980
DB 3296 GAGAGTCTACAGAAAGATGATGATGCTTTCAGGTTCTCAAGAAATTTGAAGAGAG 3355
QY 981 PheLysGlnProIleProAlaSerSerMetProGlnGlnIleGlyIleProThrSerSer 1000
DB 3356 TTCAAAACAAACCAATCTGCTCTTCCATGCAAGCAAGCAAAATGCAATCTCAACAGTCT 3415
QY 1001 LeuThrGlnValAlaHisSerAlaGlyArgArgPheIleValSerProValProGluSer 1020
DB 3416 TTAACCTCAAGTTGTTCACTTTCGCGGAAGCGGTTTATAGAGTCCGTGCCAGAAAGC 3475
QY 1021 ArgLeuArgGluSerLysValPheProSerGluIleThrAspThrValAlaAlaSerThr 1040
DB 3476 CGATTAGAGAAATCAAAAGTTTCCCGAGTAAATTAACAGATACAGTGTGCCCTTCA 3535
QY 1041 AlaGlnSerProGlyMetAsnLeuSerHisSerAlaSerSerLeuSerLeuGlnGlnAla 1060
DB 3536 GCTCAGAGCCCTGGAAATGAACTTGTCTCACTGTGATCATCTTATGCTACAAAGGCC 3595
QY 1061 PheSerGluLeuArgArgAlaGlnMetThrGluGlyProAsnThrAlaProProAsnPhe 1080
DB 3596 TTTTCTGAATTCAGTACGTCGCCAAATGACAGAAAGCCCAATACAGCACTTCAAACTT 3655
QY 1081 SerHisThrGlyProThrPheProValValProProPheLeuSerSerIleAlaGlyVal 1100
DB 3656 AGTATACAGAGCAAAATTTCCAGTATACCTCTTCTTAAGAGATTCCTGAGATC 3715
QY 1101 ProThrThrAlaAlaAlaThrAlaProValProAlaThrSerSerProProAsnAspIle 1120
DB 3716 CCACACAGCAGCAGCCACACAGACAGCTCTGCAACAGAGGCCCTCTTAATATGATTC 3775
QY 1121 SerThrSerValIleGlnSerGluValThrValProThrGluGluGlyIleAlaGlyVal 1140
DB 3776 TCCACATCAAGTATTCAGTCTAGTTACAGTCCACAGAGAGGAGTCTCGAGTT 3835
QY 1141 AlaThrSerThrGlyValValThrSerGlyLysLeuProIleProProValSerGluSer 1160
DB 3836 GCCACAGCAGAGGATGTGTAATTCAGGTGTCTCCCATACCACTGTGTCTGATATA 3895
QY 1161 ProValLeuSerSerValValSerSerIleThrIleProAlaValValSerIleSerThr 1180

Db 3896 CCAGTACTTCCAGGCTAGTTCAAGTATCACAAATACCTGAGTTGCTCAATATCTACT 3955
QY 1181 ThrSerProSerLeuGlnValProThrSerThrSerGluIleValValSerSerThrala 1200
Db 3956 AATCCCCCTGCTCACTTCAAGTCCCAATCCACATCCATCGATGCTGTTCTCTAGTACAGA 4015
QY 1201 LeuIYrProSerValThrValSerAlaThrSerAlaSerAlaGlyGlySerThrAlaThr 1220
Db 4016 CTGTATCTTCTGTAACAGTTCACAGTTCAGCACTTCAAGCTTGCAGGGGGGAGTACTCTACC 4075
QY 1221 ProGlyProIYsProProAlaValValSerGlnAlaAlaGlySerThrThrValGly 1240
Db 4076 CCAGGTCCCAAGCTCCAGCTTAGTATCTGACGACGACGACGACGACGACGACGACGACG 4135
QY 1241 AlaThrLeuThrSerValSerThrThrThrSerPheProSerThrAlaSerGlnLeuSer 1260
Db 4136 GCACATTAAACATCAAGTTCACCACTTCAATCCCAAGACAGCTTCAAGCTGACC 4195
QY 1261 IleGlnLeuSerSerSerThrSerThrProThrLeuAlaGluThrValValValSerAla 1280
Db 4196 ATTCAAGCTTAGCAGAGTACTTCTACTCTTCACTTACGTAAGGAAACCGTGTAGTACGCCA 4255
QY 1281 HisSerLeuAspIYrThrSerHisSerSerThrThrGlyLeuAlaPheSerLeuSerAla 1300
Db 4256 CACTCAGTACAGTAAAGACATCTCATAGACAGTACCACTGATGGCTTCTCCCTCTGCA 4315
QY 1301 ProSerSerSerSerSerProGlyValAlaGlyValSerSerIYrIleSerGlnProGlyGly 1320
Db 4316 CCACTCTCTCTCTCTCTCTCTGAGACAGAGTGTCTAGTATATTTCTTCAAGCTGTGGGG 4375
QY 1321 LeuHisProLeuValIleProSerValIleAlaSerThrProIleLeuProGlnAlaAla 1340
Db 4376 CTGATCTCTTGGTGCATTTCCATCGAGTACGATAGCTTCACTCTATTTCTTCCCAAGCAGCA 4435
QY 1341 GlyProThrSerThrProLeuLeuProGlnValProSerIleProProLeuValGlnPro 1360
Db 4436 GGAACCTACTTCAACCTTATTACCCCAAGTACTGATCCCACTTGGTGAAGACCT 4495
QY 1361 ValAlaAsnValProAlaValGlnGlnThrLeuIleHisSerGlnProGlnProAlaLeu 1380
Db 4496 GTTGCATATGATGCTGTGTGACAGACACATTAATCATAGTACGCTTCAACCAAGCTTGG 4555
QY 1381 LeuProAsnGlnProHisThrHisCysProGluValAlaAspSerAspThrGlnProIYsAla 1400
Db 4556 CTTCACCAACGACCCCACTACTCTCATTTGCTGAAAGTATCTGATACCAACCAAGCTTGA 4615
QY 1401 ProGlyIleAspAspIleIYrThrLeuGlnGluIleIYrLeuArgSerLeuPheSerGlnHis 1420
Db 4616 CCTGGAATGATGACATAAAGACTCTGAAGAAAGGCTGGCTCTCTTCAAGTGAACAC 4675
QY 1421 SerSerSerGlyValaGlnHisAlaSerValSerLeuGlnThrSerLeuValIleGlnSer 1440
Db 4676 AGCTCATCTGGAGCTCAGCATGCTGCTCTCACTGAGACCTCACTAGTACATAGAGAGC 4735
QY 1441 ThrValThrProGlyIleProThrThrAlaValAlaProSerIYsLeuLeuThrSerThr 1460
Db 4736 ACTGTACACAGGACATCCCACTACTGCTGTGACCAAGAAACTCTGACTTCTTACCC 4795
QY 1461 ThrSerThrCysLeuProProThrThrAsnLeuProLeuGlyThrValAlaLeuProValThr 1480
Db 4796 ACAAGTACTTGTACTTACCAACCAATTTACCACTAGGAAAGCTTCTTGGCAAGTTTAA 4855
QY 1481 ProValValThrProGlyGlnValSerThrProValSerThrThrThrSerGlyValIYs 1500
Db 4856 CCAGGTGTACACCTGGGGCAAGTTCACCCAGTACAGACTACATCAAGAGTGAAGAA 4915
QY 1501 ProGlyThrAlaProSerIYsProProLeuThrIYsAlaProValLeuProValGlyThr 1520
Db 4916 CCGTGAAGCTGCTCCCTCCAAAGCAGCTTACTTAAGGCTCCGATGTGTGCAATGGGTACT 4975
QY 1521 GluLeuProAlaGlyThrLeuProSerGlnGlnLeuProProPheProGlyProSerLeu 1540
Db 4976 GAACTTCCAGCAGTACTCTTCAACGACGAGCAGCTGCACTTTTCCAGGACCTTCTCTTA 5035
QY 1541 ThrGlnSerGlnGlnProLeuGlnAspLeuAspAlaGlnLeuArgThrThrLeuSerPro 1560
Db 5036 ACCCAGTCCCAACCACTCTAGAGGATCTTGAATGCTCAATTAGAAAGAACCACTAGTCA 5095
QY 1561 Glu***IleThrValThrSerAlaValGlyProValSerMetAlaAlaProThrAlaIle 1580
Db 5096 GAGATGATCAGAGTACTTGTGGCTGTGTCTGTGTCCATGCGGAGCTTCCAAACAGCAATC 5155
QY 1581 ThrGluAlaGlyThrGlnProGlnIYsGlyValSerGlnValIYsGluGlyProValLeu 1600
Db 5156 ACAGAAAGAGAAACACAGCTCAGAAAGGTGTCTTCAAGTCAAAAGAGCCCTTCTTA 5215
QY 1601 AlaThrSerSerGlyValaGlyValPheIYsMetGlyIYrPheGlnValSerValAlaAla 1620
Db 5216 GCACTACTTACAGAGCTGGTGTGTTTAAAGTGAAGGACATTTCAAGTTCTGTGTGACAGA 5275
QY 1621 AspGlyAlaGlnIYsGluGlyIYsAsnIYsSerGlnAspAlaIYsSerValHisPheGlu 1640
Db 5276 GACGTGCCCCGAAAGAGGCTAAAAATAGTCAAGAAAGTGAAGTGTTCATTGTGAA 5335
QY 1641 SerSerThrSerGlnSerSerValLeuSerSerSerSerProGlnSerThrLeuValIYs 1660
Db 5336 TCCAGCACTCAGAGTCTCTCAGTCTATCAAGTGTAGTCCAGAGAGTACTTGTGTAA 5395
QY 1661 ProGlnProAsnGlyIleThrIleProGlyIleSerSerAspValProGlnSerAlaHis 1680
Db 5396 CCAGAGCCCAATGACATPACCATCTCTGATATCTTCAAGTGTCCAGAGAGTGCAC 5455
QY 1681 IYsThrThrAlaSerGlnAlaIYsSerAspThrGlyGlnProThrIYsValGlyArgPhe 1700
Db 5456 AAAAATCTGCTCTCAGAGGCAAAAGTCAAGACATGGAGCTTACCAAGTGTGACGTTT 5515
QY 1701 GlnValThrThrThrAlaAsnIYsValGlyIYrArgPheSerValSerIYrThrGluAspIYs 1720
Db 5516 CAGGTGCACTACAGCAACAAAGTGGTGTCTTCTCTATCAAAATCTGAGAGCAG 5575
QY 1721 IleThrAspThrIYsIYsGluGlyProValAlaSerProProPheMetAspLeuGln 1740
Db 5576 ATCACTGACCAAAAGAAAGAGGACAGTGGCATCTCTCTTTATGATTTGAAACAA 5635
QY 1741 AlaValLeuProAlaValIleProIYsGluIYsProGlnLeuSerGlnProSerHis 1760
Db 5636 GCTGTTCTTCTGCTGTGATACCAAGAAAGAAAGACCTGAACTGTCAAGGCTTCAAT 5695
QY 1761 LeuAsnGlyProSerSerAspProGlnAlaAlaPheLeuSerArgAspValAspAspGly 1780
Db 5696 CTAAATGGGCGCTCTCTGACCCGAGGCGGCTTTTAAAGTAAAGGATGTGATGTGT 5755
QY 1781 SerGlySerProHisSerProHisGlnLeuSerSerIYsSerLeuProSerGlnAsnLeu 1800
Db 5756 TCCGATGATCCACACTGCGCCCATCATGAGTCACTCAAAAGCCCTTCTGACCAATCTA 5815
QY 1801 SerGlnSerLeuSerAsnSerPheAsnSerSerIYrMetSerSerAspAsnGlnSerAsp 1820
Db 5816 AGTCAAGCCTTAGTATATTCACTTAACTCTTACATAGTAGAGCAATAGTAGT 5875
QY 1821 IleGluAspGluAspLeuIYsLeuGlnLeuArgArgLeuArgAspIYsHisLeuIYsGlu 1840
Db 5876 ATCGAAGATGAAGACTTAAAGTTAGAGTGCAGACATCGAGATPAACATCTCAAGAG 5935
QY 1841 IleGlnAspLeuGlnSerArgGlnIYsHisGluIleGlnSerLeuIYrThrIYsLeuGly 1860
Db 5936 ATTTCAGACCTCAGAGTGCAGAGCAAGTGAATTTGATCTTTGTATACCAATCGGCG 5995
QY 1861 IYsValProProAlaValIleIleProProAlaAlaProLeuSerGlyArgArgArgArg 1880
Db 5996 AAGGTGCCCTCGCTTATTATTTCCCAAGTGTCTCCCTTTGAGGAGAAAGAGAGCA 6055
QY 1881 ProThrIYsSerIYsGlySerIYsSerSerArgSerSerSerLeuGlyAsnIYsSerPro 1900
Db 6056 CCACTAAAGCAAGGACAGCAAAATCTAGTGAAGCACTTCTTGGGGAATPAAGAGCC 6115

QY 1901 GlnLeuSerGlyAsnLeuSerGlyGlnSerAlaIleSerValLeuHisProGlnGlnThr 1920
 DB 6116 CAGCTTCAAGTAACCTGCTGGTCAAGATGACAGCTTCAAGTCTTGACACCCCAAGACAGC 6175
 QY 1921 LeuHisProGlnGlyAsnLeuSerGlyGlnSerAlaIleSerValLeuHisProGlnGlnThr 1940
 DB 6176 CTCACCTTCTGGCAACATCCAGAGTCCGAGCAATACGCTTTACAGCCCTTAAAG 6235
 QY 1941 ProSerProSerSerAspAsnLeuTySerAlaPheThrSerAspGlyAlaIleSerVal 1960
 DB 6236 CCACTCTCCCTCCAGTGAACACCTCTATTCAGCTTCAACAGTATGATGCTGCTCAGTA 6295
 QY 1961 ProSerLeuSerAlaProGlyGlnGlyThrSerSerThrAsnThrValGlyAlaThrVal 1980
 DB 6296 CCAAGCTTCTCTCCAGGTCAAGAACCAAGACCAACCACTGTTGGGGCAACAGTG 6355
 QY 1981 AsnSerGlnAlaIleAspAlaGlnProProAlaMetThrSerSerArglyGlyThrPhe 2000
 DB 6356 AACAGCCAGCCGCGCCAGCTCAAGCTCCTGCTCCATGACGTCACAGCAAGAGGCAACATTC 6415
 QY 2001 ThrAspAspLeuHisIleValAspAsnTyrAlaArgAspAlaMetAsnLeuSerGly 2020
 DB 6416 ACAGATGACTTGAACAAGTGTGTGACAAATTGGGCGCCAGATGCAATGATCTCTCAGGC 6475
 QY 2021 ArgArgGlySerIleGlyHisMetAsnTyrGlyGlyProGlyMetAlaArgIlePheSer 2040
 DB 6476 AGGAGAGAGAGCAAGGAGCAATATATGAGGGCCCTGGAATGCAAGAGATCTCT 6535
 QY 2041 AlaProGlyGlnLeuGlyIleSerMetThrSerAsnLeuGlyIleSerAlaProIleSer 2060
 DB 6536 GCACCTGGGCACTGTGATTCATTCATGACCTCGAACCCTGGGTGCTCTGCCCATCTCT 6595
 QY 2061 AlaIleSerAlaThrSerLeuGlyHisPheThrIleSerMetCysProProGlnGlnThr 2080
 DB 6596 GCACATCAAGTACTCTCTCTAGTCACTTCAACAGTCTATGTGCTCCCGCCCAACAGAT 6655
 QY 2081 GlyPheProAlaThrProPheGlyAlaGlnTyrSerGlyThrGlyProAlaProGln 2100
 DB 6656 GGCCTTCCAGCTACCCCACTTGGGCTCAATGAGTGGAGCGGTGGCCCAAGCAACAG 6715
 QY 2101 ProLeuGlyGlnPheGlnProValGlyThrAlaSerLeuGlnAsnPheAsnIleSerAsn 2120
 DB 6716 CCACCTGGCAGTCTCAACCTGTGGAACTGCTCTCTGCAAGATTTCAACATCAAGAT 6775
 QY 2121 LeuGlnIleSerIleSerAsnProProGlySerAsnLeuArgThr 2136
 DB 6776 TTGCAAGATCCATCAGCAACCCCAAGCTCAACCTGCGGACCACT 6823

RESULT 3
 AR302591 6690 bp DNA linear PAT 12-JUN-2003
 LOCUS AR302591
 DEFINITION Sequence 7 from patent US 6541252.
 ACCESSION AR302591
 VERSION AR302591.1 GI:31690889
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 6690)
 AUTHORS Walke, D.W., Hilbun, E., Donoho, G. and Turner, C.A. Jr.
 TITLE Human kinases and polynucleotides encoding the same
 JOURNAL Patent: US 6541252-A 7 01-APR-2003;
 FEATURES
 source 1..6690
 /organism="Unknown"
 /mol_type="genomic DNA"

ALIGNMENT SCORES:

Pred. No.: 1.43e-256 Length: 6690
 Score: 10753.50 Matches: 2136
 Percent Similarity: 95.83% Conservative: 0
 Best Local Similarity: 95.83% Mismatches: 0

Query Match: 99.46% Indels: 93
 DB: 6 Gaps: 1
 US-10-010-720-14 (1-2136) x AR302591 (1-6690)

QY 1 MetSerGlyGlyAlaIleAspAlaGlnProProAlaMetThrSerSerArglyGlyThrPheLeuSerPro 20
 DB 1 ATGTCTGGCGGCGCCAGAGAAAGCAGACAGACATCCCGGTCTCTGTTCTCTGCGCG 60
 QY 21 ProAlaProAlaProIleAsnGlySerSerSerAspSerSerValGlyGlyLeuGly 40
 DB 61 CCGGCTCTGCGCCCAAGATAGCTCCAGTCTCGATTCTCTCGTGGGAGAAACTGGGA 120
 QY 41 AlaAlaAlaIleAspAlaValThrGlyArgThrGlnGlyTyrArgArgArgGlnSer 60
 DB 121 GCCCGCGCGCCGACCTGTGACCGGACGAGCAAGAGTACAGGCGCGCCCACT 180
 QY 61 MetAspIleAspSerArglyAlaIleAlaIleThrThrThrThrThrGlnHisArgPhePhe 80
 DB 181 ATGACAAAGGACAGCCGTGGGCGCGCGGACCACTACCACTAGACACCCCTTCTTC 240
 QY 81 ArgArgSerValIleCysAspSerAsnAlaThrAlaLeuGlnLeuProGlyLeuProLeu 100
 DB 241 CGCCGAGCGTCACTGCGACTCCATGCACTGCACTGAGCTTCCCGGCTCTCTCT 300
 QY 101 SerLeuProGlnProSerIleProAlaAlaValProGlnSerAlaProProGlnProHis 120
 DB 301 TCCCTGCGCCAGCCAGCACTCCCGGGCTGCTCCCGCAGATGCTCCACCGAGCCCGAC 360
 QY 121 ArgGlnGlnThrValThrAlaThrAlaThrSerGlnValAlaGlnGlnProProAlaAla 140
 DB 361 CGGGAAGAGACCGGACCGCCAGCCGCACTTCCACAGGTACCCAGCGCTCCAGCGCT 420
 QY 141 AlaAlaProGlyGlnGlnAlaValAlaGlyProAlaProSerThrValProSerSerThr 160
 DB 421 GCCGCGCTGGGAGAACGCGCTGCGGCGGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
 QY 161 SerIleAspArgProValSerGlnProSerLeuValGlySerIleGlyGlnProProPro 180
 DB 481 AGCAAAACCCCGCCAGATGCCAGCTTACCTTGTGGGAGCAAAAGAGAGAGCCCGCG 540
 QY 181 AlaArgSerGlySerGlyGlySerAlaIleGlnProGlnGlnGlnGlnGlnGlnGln 200
 DB 541 GCGAAGATGGCAGCGCGCGCGCGCCAGGACCAAGAGCCAGGAGGAGGAGGAGGAGGAG 600
 QY 201 GlnAspAspIleGlnGlnLeuGlnThrIleValAlaGlyMetSerAsnAspGlyArgPhe 220
 DB 601 CAGGATGATATCGAAGAGCTGAGACCAAGCCCTGGGAATGCTCAACGATGGCGCTTT 660
 QY 221 LeuIlePheAspIleGlnIleGlyArgGlySerPheIleThrValTyrIleGlyLeuAsp 240
 DB 661 CTCAAGTTTGAATCGAATACGGCAGAGGCTCTTAAAGCGTCTTCAAAAGGTCTGGAC 720
 QY 241 ThrGlnThrThrValGlnValAlaTyrPheGlnLeuGlnAsnArgIleThrIleSer 260
 DB 721 ACTGAACACCGTGGAGTGGCTGTGTGAATCGAGATCGAAATTTAAACAAAGTCT 780
 QY 261 GlnArgGlnArgPheIleGlnGlnAlaGlnLeuMetLeuIleGlnGlnHisProAsnIle 280
 DB 781 GAGAGGCGAGATTTAAAGAGAGCTGAATGTTAAAGAGCTTCAGACATCCCAATAT 840
 QY 281 ValArgPheIleArgSerTyrPheIleThrValIleGlyGlyGlyGlyValIleVal 300
 DB 841 GTTAATTTTAAATTCCTCGGAAATCCACAGTAAAGGAAAGGAAAGTGTCTTTTGGTG 900
 QY 301 ThrGlnLeuMetThrSerGlyThrIleIleIleIleIleIleIleIleIleIleIle 320
 DB 901 ACTGAATTTATGACGCTGCGAACACTTAAAGGATCTGAAAGGTTTAAAGTATGAG 960
 QY 321 IleIleValLeuArgSerTyrCysArgGlnIleLeuGlyLeuGlnPheLeuHisThr 340
 DB 961 ATCAAAAGTCTAAAGAGCTGTGTGCTCGATCTTAAAGTCTTCAAGTTCTTCAATAC 1020


```
QY 341 ArgThrProProIleIleHisArgAspLeuLysCysAspAsnIlePheIleThrGlyPro 360
Db 1021 CGAAGTCCAGCCTATCATTCACCGCAATCTTAAATGAGACACATCTTATACCGGCGCT 1080
QY 361 ThrGlySerValIleGlyAspLeuGlyLeuAlaThrLeuLysArgAlaSerPheAla 380
Db 1081 ACTGGCTCAGTCAAGATTGGAGACCTCGGTCTGGCAGCCCTGAGACGGGCTTCTTTGGCC 1140
QY 381 LysSerValIleGlyThrProGluPheMetAlaProGluMetTyrGluGluLysTyrAsp 400
Db 1141 AAGAGTGTGATGAGTACCCAGAGTTCATGCGCCCTCGAGATGTATGAGAGAAATATGAT 1200
QY 401 GluSerValAspValTyrAlaPheGlyMetCysMetLeuGluMetAlaThrSerGluTyr 420
Db 1201 GAATCCGTGAGCTTATGCTTTTGGAGTGTGAGATGCTTGAGATGGCTATCTCAATAT 1260
QY 421 ProTyrSerGluCysGlnAsnAlaIleGluIleTyrArgValThrSerGlyValLys 440
Db 1261 CCTTACTCGAGTGGCAAAATGCTGCAAGATCTTACCTCGGTGACCAATGGGGTGAAG 1320
QY 441 ProLysSerPheAspLysValAlaIleProGluValLysGluIleIleGluGlyCysIle 460
Db 1321 CCAGCAGATTGTGACAAAGTACATCTGAGAGTGAAGAAATATATTGAGAGATGATA 1380
QY 461 ArgGlnAsnLysAspGluArgTyrSerIleLysAspLeuLeuAsnIleAlaPhePheGln 480
Db 1381 CGAAGAAACAAAGATGAAAGATTCATCCATCAAGACCTTTGAAACCAATGCTTTCCAA 1440
QY 481 GluGluThrGlyValArgValGluLeuAlaGluGluLysAspAspGlyGluLysIleAlaIle 500
Db 1441 GAGGAAACAGAGATACGGGTAGATTTAGCAGAGAAAGATGATGAGAAAGAAATATGCA 1500
QY 501 LysLeuThrLeuAlaGlyIleGluAspIleLysLysLeuLysGlyLysTyrLysAspAsnGlu 520
Db 1501 AAATATAGGCTACGATATTGAAGATATTAGAAATTTAAAGGAAATTAAGAAATTAAGAA 1560
QY 521 AlaIleGluPheSerPheAspLeuGluArgAspValProGluAspValAlaGlnGluMet 540
Db 1561 GCTTTTATGATTTCTTTTGAATTAGAGAGAGATCTCCAGAAAGATGTTGACAAAGAAATG 1620
QY 541 ValGluSerGlyTyrValCysGluGlyAspHisIleLysThrMetAlaLysAlaIleLysAsp 560
Db 1621 GTAGAGTCTGGATATGCTGTGAAGGTGATCAAGACCATGGCTTAAAGCTTATCAAAAGAC 1680
QY 561 ArgValSerLeuIleLysArgLysArgGluArgGluArgGluLeuValArgGluGluGln 580
Db 1681 AGAGTATCATTAATTAAAGAGAAACGAGACGCGGAGTTGGTACGGGAGGAGAAAGAA 1740
QY 581 LysLysLysGlnGluGluSerLeuLysGlnGlnValGluGluSerValAspGln 600
Db 1741 AAAAAAAGCAGAGAGAGAGACAGTCTCAACAGCAGGTAGAACATCCAGTGTCTCCAG 1800
QY 601 ThrGlyIleLysGlnLeuProSerAlaSerThrGlyIleProThrAlaSerThrThrSer 620
Db 1801 ACAGGATCAAGCAGCTCTTCTGTAGACCGGCACTTACTGCTTCTTACCACTTCA 1860
QY 621 AlaSerValSerThrGlnValGluProGluGluProGluAlaAspGlnHisGlnGlnLeu 640
Db 1861 GCTTCAGTTTCTACACAGATGAACTGAAGAACTGAGGACAGATCAACATCAACAACTA 1920
QY 641 GlnTyrGlnGlnProSerIleSerValLeuSerAspGlyThrValAspSerGlyGlnGly 660
Db 1921 CAGTACCAAGAACCCAGTATATCTGTTATCTGATGGGACGGTTCACAGTGTCTCGGGA 1980
QY 661 SerSerValPheThrGluSerArgValSerSerGlnGlnThrValSerTyrGlySerGln 680
Db 1981 TCCCTGCTCTTACAGAAATCTGAGTGAAGCAGCAGCAACACAGTTTCAATGCTTCCAA 2040
QY 681 HisGluGlnAlaHisSerThrGlyThrValProGlyHisIleProSerThrValGlnAla 700
Db 2041 CAGTACCAAGCAGCAGTCTTCAAGGACAGTCCAGGAGCATATATCTTCTACTGCAAGCA 2100
QY 701 GlnSerGlnProHisGlyValTyrProProSerSerValAlaGlnGlyGlnSerGlnGly 720
Db 2101 CAGTCTCAGCCCAATGGGATATATCCACCTTCAATGTGGCAGCAGGGGAGCAGCCAGGCT 2160
QY 721 GlnProSerSerSerSerLeuThrGlyValSerSerSerGlnProIleGlnHisProGln 740
Db 2161 CAGCCATCTCAAGTACCTTAAACAGGGGTTTCATCTTCCCAACCATTAACATCTCTGAG 2220
QY 741 GlnGlnGlnGlyIleGlnGlnThrAlaProProGlnGlnThrValGlnTyrSerLeuSer 760
Db 2221 CAGCAGCAGGAGAAATACAGACAGCAGCCCTCTCAACAGACAGTACAGTATTCATCTTCA 2280
QY 761 GlnThrSerThrSerSerGluAlaThrThrAlaGlnProValSerGlnProGluAlaPro 780
Db 2281 CAGCATCAACCTCCAGTGAAGCCACTACTGCAACAGCCAGTACAGTCAAGCTTCA 2340
QY 781 GlnValLeuProGlnValSerAlaGlyLysGln 791
Db 2341 CAACTTGTGCTCAAGTATCAGTGAAGAAACAGGGGCTTCCCACTCGACTCCACACAGAG 2400
QY 791 791
Db 2401 TACCCAGAGATTCAAATATTTGCTCTCTTCCAAAGTGGCTTCTGTTGATCCATCT 2460
QY 791 791
Db 2461 ACAAGTCTTACCCCTCCCATGGCAGCAGAGTACTGCTACCTGGAGTACTTCCACA 2520
QY 791 791
Db 2521 GTGTGAGAGCTTATGTGATCAATCTTTTATGTTCTTATGGGTGTGTAGAGAGACAG 2580
QY 791 791
Db 2581 GTTCAAGTGTCCCAAGCAGAGGAGTTAGCAAAAGCCCACTACATCTCCAGCAAA 2640
QY 792 792
Db 2641 GCAGTTTGGAGAGTACTCAGGAGTCTCTCAGGTGCTCTCTGACAGAGCCAGTGGACGA 2700
QY 808 AlaGlnProGlnAlaThrGlnProThrThrLeuAlaSerSerValAspSerAlaHisSer 827
Db 2701 GCACAGCCCAAGCTACCCAGCCGACCACTTGGCTTCTCTGTAGACATGTGACATTTCA 2760
QY 828 AspValAlaSerGlyMetSerAspGlyAsnGluAsnValProSerSerSerGlyArgHis 847
Db 2761 GATGTTGCTTCAAGATATGATGATGCAATGAAAGCTCCCATCTTCCAGTGAAGGAT 2820
QY 848 GluGlyArgThrThrLysArgHisIleLysSerValArgSerArgSerArgHisGlu 867
Db 2821 GAAGGAAGAACTACAAAGCGCATTTACGAAATCTGTAAAGAGTGGCTCTGACATGAA 2880
QY 868 LysThrSerArgProLysLeuArgIleLeuAsnValSerAsnLysGlyAspArgValVal 887
Db 2881 AAAAATCTACGCCCAAAATTAAGAAATTTGAATGTTTCAATTAAGAGACCGAGTAA 2940
QY 888 GluCysGlnLeuGlnThrHisAsnArgLysMetValThrPheLysPheAspLeuAspGly 907
Db 2941 GAATGCTCAATTGAGACTATATATGAAAAAGTGAATCACTTCAAAATTTAAGTAACTGAT 3000
QY 908 AspAsnProGluGluIleAlaThrIleMetValAsnAspAspPheIleLeuAlaIleGlu 927
Db 3001 GACAAACCCGAGAGATGACCAACATTAATGAGCAACATGATTTATTTAGCAATAGAG 3060
QY 928 ArgGluSerPheValAspGlnValArgGluIleIleGluLysAlaAspGluMetLeuSer 947
Db 3061 AAGAGTGTGTTGTGATCAAGTGTGAGAAATTAATTAAGAAAGTGAATGCTCACT 3120
QY 948 GluAspValSerValGluProGluLysAspGlnGlyLeuGlnSerLeuGlnGlyLysAsp 967
Db 3121 GAGGATGTCAAGTGTGAACCAAGGAGTATCAAGGATTTGAGAGTCTCAAGAAAGAGAT 3180
QY 968 AspTyrGlyPheSerGlySerGlnLysLeuGluGluGluPheLysGlnProIleProAla 987
```

Db 3181 GACTATGCTTTTCAGGTTCTCAGAAATTGGAAGAGAGTTCAAACCACTTCCTGG 3240
| | | | |
Qy SerSerMetProGlnGlnIleGlyIleProThrSerSerLeuThrGlnValAlaHisSer 1007
| | | | |
Db 3241 TCTTCATGCGACAGCAAAATAGGCAATTCCTACCAAGTTCTTTAACTCAAGTTGTTCACTTC 3300
| | | | |
Qy 1008 AlaGlyArgArgPheIleValSerProValProGlnSerArgLeuArgGlnSerVal 1027
| | | | |
Db 3301 GCGGGAAGGGGTTTATGAGATGCTCGGCCAGAAAGCGGATATACAGAAATCAAAAGTT 3360
| | | | |
Qy 1028 PheProSerGlnIleThrPheThrValAlaAlaSerThrAlaGlnSerProGlyMetAsn 1047
| | | | |
Db 3361 TTCCCAAGTAAATACAGATACAGTTGCTGCTCTACAGCTCAGAGCCCTGGAAATGAAC 3420
| | | | |
Qy 1048 LeuSerHisSerAlaSerSerLeuSerLeuGlnAlaPheSerGlnLeuAlaArgAla 1067
| | | | |
Db 3421 TTGTCTCAGCTGCATCATCCCTTAGCTCAACAGAGCCCTTTCTGAACTTAGACGTGCC 3480
| | | | |
Qy 1068 GlnMetThrGlnGlyProAlaThrAlaProProAsnPheSerHisThrGlyProThrPhe 1087
| | | | |
Db 3481 CAATATGACAAAGAACCAAAVACAGACCTCCAAACTTATGTCATACAGAACCAACTTT 3540
| | | | |
Qy 1088 ProValValProProPheLeuSerSerIleAlaGlyValProThrThrAlaAlaAlaThr 1107
| | | | |
Db 3541 CCAAGTACTACTCTCTTTCTTAAGTACATGCTGAGAGTCCCAACCAAGCAGAGCCACA 3600
| | | | |
Qy 1108 AlaProValProAlaThrSerSerProProAsnAspIleSerThrSerValIleGlnSer 1127
| | | | |
Db 3601 GCACCACTCCCTCGCAACAGACGCCCTCTATATGACATTCACATCAATGATATCACTCT 3660
| | | | |
Qy 1128 GlnValThrValProThrGlnGlnGlyIleAlaGlyValAlaThrSerThrGlyValVal 1147
| | | | |
Db 3661 GAGGTTCAGTGGCCCACTGAAGAGGGATGCTGAGTTGCCACAGCAGACAGGTGCTGA 3720
| | | | |
Qy 1148 ThrSerGlyGlyLeuProIleProProValSerGlnSerProValLeuSerSerValVal 1167
| | | | |
Db 3721 ACTTCAGATGCTCTCCCATATCAACCTGTGTCTGATACCAAGTACTTCCAGGTAGTT 3780
| | | | |
Qy 1168 SerSerIleThrIleProAlaValAlaSerIleSerThrThrSerProSerLeuGlnVal 1187
| | | | |
Db 3781 TCAGATATCACAAAATCTGAGTGTGTCTCAATATCTACTACATCCCTCCCTCACTCAAGTC 3840
| | | | |
Qy 1188 ProThrSerThrSerGlnIleValValSerSerThrAlaLeuArgProSerValThrVal 1207
| | | | |
Db 3841 CCCACATCCACATCTGAGATGTTGTTCTAGTACAGCATGTTATCTCTTCCAGTACAGTT 3900
| | | | |
Qy 1208 SerAlaThrSerAlaSerAlaGlyGlySerThrAlaThrProGlyProGlyProProAla 1227
| | | | |
Db 3901 TCAGCACTTACAGCTCTGCAAGGGGAGTACTGCTACCCAGAGTCCTAAGCTTCAGCT 3960
| | | | |
Qy 1228 ValValSerGlnGlnAlaAlaGlySerThrThrValAlaAlaThrLeuThrSerValSer 1247
| | | | |
Db 3961 GTAGATCTCAGCGAGCAGCAGCAGCACTACTGAGAGGCCAATTAACAATCAATTTCT 4020
| | | | |
Qy 1248 ThrThrThrSerPhePheProSerThrAlaSerGlnLeuSerIleGlnLeuSerSerSerThr 1267
| | | | |
Db 4021 ACCACCACTTCACTCCAGACCACTTCAAGCTGCTTCATTCAGCTTACAGCAGTACT 4080
| | | | |
Qy 1268 SerThrProThrLeuAlaGlnThrValValAlaSerAlaHisSerLeuAspIleThrSer 1287
| | | | |
Db 4081 TCTACTCTCACTTATGCGAAACCGTGGTATGAGTACGACACTCACTAGATAGACATCT 4140
| | | | |
Qy 1288 HisSerSerThrThrGlyLeuAlaPheSerLeuSerAlaProSerSerSerSerSerPro 1307
| | | | |
Db 4141 CATACAGACAAACGATGCTTCTCCCTCTCTGCAACCATCTTCTCTCTCTCTCT 4200
| | | | |
Qy 1308 GlyAlaGlyValSerSerThrIleSerGlnProGlyGlyLeuHisProLeuValIlePro 1327
| | | | |
Db 4201 GAGGAGGAGGTCTAGATTATTTCTTCAGCTGTGGCTGCATCTTGGTATCTTCA 4260
| | | | |
Qy 1328 SerValIleAlaSerThrProIleuProGlnAlaAlaGlyProThrSerThrProLeu 1347
| | | | |
Db 4261 TCAGTGAATGCTTCACTCTATTTCTTCCCAAGCAGCAGACCTAATCTTCACTTCA 4320
| | | | |

Qy 1348 LeuProGlnValProSerIleProProLeuValGlnProValAlaAsnValProAlaVal 1367
| | | | |
Db 4321 TTACCCCAAGTACTAGTATCCCACTTGTGACAGCTGTTGCCATGCTGCTGTGA 4380
| | | | |
Qy 1368 GlnGlnThrLeuIleHisSerGlnProGlnProAlaLeuLeuProAsnGlnProHisThr 1387
| | | | |
Db 4381 CAGCAGACACTAATTCATGATGAGCTTCAACCAAGCTTGTGCTTCCCAACCAAGCCCACTACT 4440
| | | | |
Qy 1388 HisCysProGlnValAspSerAspThrGlnProGlyAlaProGlyIleAspAspIleCys 1407
| | | | |
Db 4441 CATTGCTCGAAGTAAATTTCTGATACAAACCCAAAGCTCTGGAATTGATGACATTAAG 4500
| | | | |
Qy 1408 ThrLeuGlnGlnLeuSerLeuArgSerLeuPheSerGlnHisSerSerSerGlyAlaGlnHis 1427
| | | | |
Db 4501 ACTCTGAAAGAAAGCTCGCGTCTCTGTTCAAGAAACAGCTCACTGAGCTCAGCAT 4560
| | | | |
Qy 1428 AlaSerValSerLeuGlnThrSerLeuValIleGlnSerThrValThrProGlyIlePro 1447
| | | | |
Db 4561 GCTCTGTCTCACTGAGAGACCTCAATGATAGAGAGCACTGTCAACACAGGACATCCCA 4620
| | | | |
Qy 1448 ThrThrAlaValAlaProSerIleLeuLeuThrSerThrThrCysLeuProPro 1467
| | | | |
Db 4621 ACTACTGCTGTTCACCAAGCAAACTCTGACTTTCACCAAGTACTTCTTACCAACA 4680
| | | | |
Qy 1468 ThrAsnLeuProLeuGlyThrValAlaLeuProValThrProValValThrProGlyGln 1487
| | | | |
Db 4681 ACCAATTTACCACTAGGAACAGTTGCTTCCAGTTACACAAAGGTGCACACCTGGGCA 4740
| | | | |
Qy 1488 ValSerThrProValSerThrThrThrSerGlyValIleProGlyThrAlaProSerIle 1507
| | | | |
Db 4741 GTTTCACCCCACTGACAGCCTACTACATCAGAGGTGAACCTGGAACTGCTCTCCACAG 4800
| | | | |
Qy 1508 ProProLeuThrIleValAlaProValIleuProValGlyThrGlnLeuProAlaGlyThrLeu 1527
| | | | |
Db 4801 CCACCTTAATCTAAGGCTCCGAGTCTCCAGTGGGTACTGAATCTTCCAGCAGTACTCTA 4860
| | | | |
Qy 1528 ProSerGlnGlnLeuProProPheProGlyProSerLeuThrGlnSerGlnProLeu 1547
| | | | |
Db 4861 CCNAGCAGAGCAGTGCACCTTTCCAGGACCTTCTCAACCCAGTCCAGCAACTCTA 4920
| | | | |
Qy 1548 GlnAspLeuAspAlaGlnLeuArgArgThrLeuSerProGlu***IleThrValThrSer 1567
| | | | |
Db 4921 GAGGATCTTGAATGCTCAATGAGAAACAATGATGTCAGAGATKATACAGTACTCT 4980
| | | | |
Qy 1568 AlaValGlyProValSerMetAlaAlaProThrAlaIleThrGlnAlaGlyThrGlnPro 1587
| | | | |
Db 4981 GCGGTGGTCTGTGTCCATGCGGCTCCAAACAGCAATCAAGAAACAGCAAGCT 5040
| | | | |
Qy 1588 GlnIleGlyValSerGlnValIleGlyProValLeuAlaThrSerSerGlyValAlaGly 1607
| | | | |
Db 5041 CAGAAAGGTGTTTCTCAAGTCAAAGAGGCTCTGCTCAAGCACTTCAAGAGCTGCTGT 5100
| | | | |
Qy 1608 ValPheIleMetGlyArgPheGlnValSerValAlaAlaAspGlyAlaGlyIleGly 1627
| | | | |
Db 5101 GTTTTAAAGATGAGCGAATTCAGTTCGTGTGAGAGAGCGGTGCCAAGAAAGAGGT 5160
| | | | |
Qy 1628 LysAsnLysSerGlnAspAlaIleSerValHisPheGlnSerSerThrSerGlnSerSer 1647
| | | | |
Db 5161 AAAAATTAAGTACAGAAATGCAAGTCTGTCTTGAATCCAGCACTCAGAGTCTCA 5220
| | | | |
Qy 1648 ValLeuSerSerSerSerProGlnSerThrLeuValIleProGlnProAsnGlyIleThr 1667
| | | | |
Db 5221 GTGCTATCAAGTATGATCCAGAGAGTACTTGTGAAACCAAGACCCAAATGGCAATACC 5280
| | | | |
Qy 1668 IleProGlyIleSerSerAspValProGlnSerAlaHisIleThrThrAlaSerGlnAla 1687
| | | | |
Db 5281 ATCCCTGTATCTCTTCAGATGTGCCAAGATGTCCCAAAATCTAGCTTCAAGAGCA 5340
| | | | |
Qy 1688 LysSerAspThrGlnGlnProThrIleValGlyArgPheGlnValThrThrThrAlaAsn 1707
| | | | |
Db 5341 AAGTACAGACCTGGCAGGCTTCAAGAGTTGACGTTTTCAGGTGACAACTACAGAAAC 5400
| | | | |

QY 1708 LysValGlyArgPheSerValSerIleThrAspThrIleValGlu 1727
 DB 5401 AAAGGCGGCTGCTTCTCTGATCAAAAGTGAAGCAAGATACGACCAAAAGAAAGAA 5460
 QY 1728 GlyProValAlaSerProPheMetAspLeuGlnAlaValLeuProAlaValIle 1747
 DB 5461 GGAACAGAGGAGCTCTCTCTTTATGATTGGAAACAGCTGTTCTTCTGCTGATTA 5520
 QY 1748 ProGlyGlyGlyLeuProGlyLeuSerGluProSerHisLeuAsnGlyProSerSerAsp 1767
 DB 5521 CCMAAGAAAGAGAACCTGAACCTGACAGAGCTTCAACATCTTAATGGGCGCTCTGAC 5580
 QY 1768 ProGluAlaAlaPheLeuSerArgAspValAspAspGlySerGlySerProHisSerPro 1787
 DB 5581 CCGAGGCGGCTTTTAAAGTAGGATGTGATGTGTGTTCCGATGACACACTCGCC 5640
 QY 1788 HisGlnLeuSerSerLeuSerLeuProSerGlnAsnLeuSerGlnSerLeuSerAsnSer 1807
 DB 5641 CATCAGCTGAGCTCAAAAGAGCTTCTTACCGAATCTTAAGTCAAGGCTTAATATCA 5700
 QY 1808 PheAsnSerSerIleMetSerSerAspAsnGlySerAspIleGluAspGluAspLeuIys 1827
 DB 5701 TTTAACTCTCTACATGATAGTACGACCAATGATGATGATGAAAGACTTAAG 5760
 QY 1828 LeuGluLeuArgArgLeuArgAspIleHisLeuIysGluIleGlnAspLeuGlnSerArg 1847
 DB 5761 TTAAAGCTGCAAGCACTAGAGATTAACATCTCAAGAGATTCAGGAGCTCAGAGTCTG 5820
 QY 1848 GlnIysHisGluIleGluSerLeuIleThrIysLeuGlyIleValProProAlaValIle 1867
 DB 5821 CAGAAAGCATGAATTTGATCTTTGTATCCAACTGGGCAAGGAGGCTCCCTGCTGATTT 5880
 QY 1868 IleProProAlaAlaProLeuSerGlyArgArgArgProThrIysSerLeuGlySer 1887
 DB 5881 ATTCCCCAGCTGCCCTTTCAGGAGAGAGACGACGACCACTAAAGAAAGGAGC 5940
 QY 1888 LysSerSerArgSerSerSerLeuGlyAsnIysSerProGlnLeuSerGlyAsnLeuSer 1907
 DB 5941 AAATCTAGTCGAAGAGATTCTCTGGGAAATAAAGCCCCAGCTTTCAGTAACTGCTCT 6000
 QY 1908 GlyGlnSerAlaAlaSerValLeuHisProGlnGlnThrLeuHisProProGlyAsnIle 1927
 DB 6001 GGTACAGAGTGACCTTCACTTTCAGCCCCCAGAGAGACCTTCCCTCTGGGAGAACATC 6060
 QY 1928 ProGlnSerGlyGlnAsnGlnLeuLeuGlnProLeuIysProSerProSerSerAspAsn 1947
 DB 6061 CCAGAGTCCGGGCGAGATCACTGTTACAGCCCTTAAGCCATCTCCCTCAAGAGAAC 6120
 QY 1948 LeuIysSerAlaPheThrSerAspGlyAlaIleSerValProSerLeuSerAlaProGly 1967
 DB 6121 CTCTATTCAAGCTTCCACAGATGATGTGCCATTTCAGTACCAAGCTTTCGCTCCAGGT 6180
 QY 1968 GlnGlyThrSerSerThrAsnThrValGlyAlaThrValAsnSerGlnAlaGlnAla 1987
 DB 6181 CAAGAGACCGAGCAACAACACTGTTGGGCAACAAGAGAACCAAGCCGCGCAAGCT 6240
 QY 1988 GlnProProAlaMetThrSerSerArgIleGlyThrPheThrAspAspLeuHisLeu 2007
 DB 6241 CAGGCTCTGCAATGACGTCCAGCAGAGAGGACCATTCACAGATGACTTGCACAAATTG 6300
 QY 2008 ValAspAsnThrAlaAlaGAspAlaMetAsnLeuSerGlyArgGlySerIleValHis 2027
 DB 6301 GTAGACAAATTGGGCGCAGATGCCATGATCTCAAGCAGAGAGAAACAAAGGAGC 6360
 QY 2028 MetAsnThrGluGlyProGlyMetAlaArgIysPheSerAlaProGlyGlnLeuVal 2047
 DB 6361 ATGATATTATAGAGGCTCGAATAGCAAGAGATTTCTGCACCTGGGCAACTGTGATC 6420
 QY 2048 SerMetThrSerAsnLeuGlyGlySerAlaProIleSerAlaAlaSerAlaThrSerLeu 2067
 DB 6421 TCCATGACCTCGAACTGGGTGGCTGCCCCCATCTCTGACAGATAGAGTACTCTCTA 6480
 QY 2068 GlyHisPheThrIysSerMetCysProProGlnGlnIleGlyPheProAlaThrProPhe 2087

DB 6481 GGTCACTTACCAAGTATGATGCCCCCAGCAGATATGCTTCCAGCTACCCCATTT 6540
 QY 2088 GlyAlaGlnThrPheSerGlyThrGlyGlyProAlaProGlnProLeuGlyGlnPheGlnPro 2107
 DB 6541 GGGCTCATGTAGAGTGGAGCGGTGGCCACACACACAGCCACTTGGCCAGTTCCAACT 6600
 QY 2108 ValGlyThrAlaSerLeuGlnAspPheAsnIleSerAsnLeuGlnIleSerIleSerAsn 2127
 DB 6601 GTGGAACTGCTCTCTTGCAGATTTTCAACATCAGACATTTTGCAAAATTCATCAGACAC 6660
 QY 2128 ProProGlySerAsnLeuArgThrThr 2136
 DB 6661 CCCCCAGGCTCCAACTCGGACCACT 6687
 RESULT 4
 AR302588
 LOCUS AR302588 7149 bp DNA linear PAT 12-JUN-2003
 DEFINITION Sequence 1 from patent US 6541252.
 ACCESSION AR302588
 VERSION AR302588.1 GI:31690866
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 7149)
 AUTHORS Walke, D.W., Hilpun, E., Donoho, G. and Turner, C.A. Jr.
 TITLE Human kinases and polynucleotides encoding the same
 JOURNAL Patent: US 6541252-A 1 01-APR-2003;
 FEATURES
 source 1..7149
 /organism="unknown"
 /mol_type="genomic DNA"
 ORIGIN
 Alignment Scores:
 Pred. No.: 1,17e-254 Length: 7149
 Score: 10677.00 Matches: 2136
 Percent Similarity: 89.67% Conservative: 0
 Best local Similarity: 89.67% Mismatches: 0
 Query Match: 98.75% Indels: 246
 DB: 6 Gaps: 1
 US-10-010-720-14 (1-2136) x AR302588 (1-7149)
 QY 1 MetSerGlyGlyAlaGluIysGlnSerSerThrProGlySerLeuPheLeuSerPro 20
 DB 1 ATGCTTGGCGGCGCCGCGAGAGAGACAGCAGCACTCCCGTTCCCTGCTCTCGCG 60
 QY 21 ProAlaProAlaProIysAsnGlySerSerSerAspSerSerValGlyGlyIleGly 40
 DB 61 CCGGCTCTTCCCGCAAGATGCTCCAGTCTCCGATTCCTCCGTGGGGGAGAAATGGGA 120
 QY 41 AlAlaAlaAlaAspAlaValThrGlyArgThrGluIleIleArgArgArgGlnIleThr 60
 DB 121 GCCGCGGCGCGCAGCGCTGTGACCGGAGAGACGAGAGATACAGGCCGCCGCCCACT 180
 QY 61 MetAspIysAspSerArgGlyAlaAlaAlaThrThrThrThrThrGluHisArgPhe 80
 DB 181 ATGACAAAGACAGCGGTGGGGCGGCGGACACACACCACTGACGACCCGCTTCTTC 240
 QY 81 ArgArgSerValIleCysAspSerAsnAlaThrAlaLeuGluLeuProGlyIleProLeu 100
 DB 241 CGCGGAGGCTCATCTGCGACTCCATTCAGATGCACTGAGCTTGAGGCTTCCGCTCTCT 300
 QY 101 SerLeuProGlnProSerIleProAlaAlaValProGlnSerAlaProProGluProHis 120
 DB 301 TCCCTGCGCCAGCCAGCATCCCGCGGCTGTCCCGAGAGTCTCCACCGAGGCCCCAC 360
 QY 121 ArgGluGluThrValThrAlaThrAlaThrSerGlnValAlaGlnGlnProProAlaAla 140
 DB 361 CCGGAAAGACCGGTGACCGCCACCGCACTTCCAGATGAGCCAGCAGGCTTCCAGCGCT 420

Db 2581 CTTCCCATCATGATGGAGCTGGCATTTACTAGCCTCTGCTCAGCTTGCTTCACTGCT 2640
 QY 791 ----- 791
 Db 2641 ACAACAGCTGCATGCCGGGGGTATCACTGTTCTTCTAGTCAGCTTCCAACTCTCTG 2700
 QY 791 ----- 791
 Db 2701 CAGCTGTGACTCAGCTGCGAAGTCAGGTTCAACCCAGCTCTCAACACAGCAGTTGAG 2760
 QY 791 ----- 791
 Db 2761 TCCATGGGAATACAGACTAACCTTGGACAGAGCTGCTGAGTTTCCACTTCTCTGGAGAT 2820
 QY 791 ----- 791
 Db 2821 GTTCTGTACAGAGGCTTCCACCTGACGTCGACACAGTACCCAGAGATTTCAATATT 2880
 QY 791 ----- 791
 Db 2881 GCTCCCTCTTCCAAAGTGGCTTCTGTTGCACTTCTACAGTCTTACCCCTCCATG 2940
 QY 791 ----- 791
 Db 2941 CCGACAGAAGTACTGCTGCTACACCTGGGTACTTTCCCAAGTGTGACGCTTATGTGAA 3000
 QY 791 ----- 791
 Db 3001 TCAATCTTTAGTCTCTATGGTGTGTAGAGAGCAGGTTCAAGTCTCCAGCCAGGA 3060
 QY 792 ----- SerThrGln 794
 Db 3061 GGGAGTTTAGCAACAGCCCCCACTACATCTCCAGCAAGCAGTTTGGAGATGACTCAG 3120
 QY 795 G1YAlSerGlnValAlaProAlaGluProValAlaValAlaGlnProGlnAlaThrGln 814
 Db 3121 GGAATCTCTCAGGTTGCTCTGTCAGAGCAGATGGAGTGAAGCAAGCCCCAAGCTTACCCG 3180
 QY 815 ProThrThrLeuAlaSerSerValAspSerAlaHisSerAspValAlaSerGlyMetSer 834
 Db 3181 CCGACCACTTGGCTCTCTGTGACAGCTGCACATTCAGATGTTGCTTCAGGTATGAGT 3240
 QY 835 AspGlyAsnGluAsnValProSerSerSerGlyArgHisGlyArgThrThrLysArg 854
 Db 3241 GATGGCAATGAGAACGTCCTCCATCTTCCAGTGAAGCATGAGAGAGAACTTACAAACGG 3300
 QY 855 HisThrArgLysSerValArgSerArgSerArgHisGlyLysThrSerArgProLysLeu 874
 Db 3301 CATTAACGAAATCTGTAAAGAGTGGCTCTGACATGAAACCTTACCGCCCAAAATTA 3360
 QY 875 ArgIleLeuAsnValSerAsnLysGlyAspArgValAlaGlyCysGlnLeuGluThrHis 894
 Db 3361 AGAATTTTGAATGTTTCAATTAAGAGAGCCGAGTAGTGAATTCATTTAGAACTCAT 3420
 QY 895 AsnArgLysMetValThrPheLysPheAspLeuAspGlyAspAsnProGluGluIleAla 914
 Db 3421 AATGAGAAATGATTACATTCMAATTTGACCTAGATGCTGCAACCCCGAGGAATAGCA 3480
 QY 915 ThrIleMetValAsnAspPheIleLeuAlaIleGluArgGlySerPheValAspGln 934
 Db 3481 ACAATTTGTGACCAATGACTTTATTTAGCAATTAAGAGAGAGGTGTTGTGATCA 3540
 QY 935 ValArgGluIleIleGluLysAlaAspGluMetLeuSerGluAspValSerValGluPro 954
 Db 3541 GTGGAGAAATTTATGAAAAAGCTGATGAATGCTCAGTGAATGTCAGTGTGAAACA 3600
 QY 955 GluGlyAspGlnGlyLeuGlyLysSerLeuGlnGlyLysAspAspTyrGlyPheSerGlySer 974
 Db 3601 GAGGCTGATCAGGATTTGAGAGAGTCAAGAGAAAGATGACTATGCTTTTACGTTCT 3660
 QY 975 GlnLysLeuGluGlyLysPheLysGlnProIleProAlaSerSerMetProGlnGlnIle 994
 Db 3661 CAGAAATTGAGAGAGAGTTCAAACCAACCAATTCCTGCTTCCATGCCACAGCAATA 3720

QY 995 GlyIleProThrSerSerLeuThrGluValAlaHisSerAlaGlyArgArgPheIleVal 1014
 Db 3721 GGCAATCTTACCAAGTCTTTTAACTCAAGATTTGTTCAATTCGCGGAGAGGGGTTTATAGTG 3780
 QY 1015 SerProValProGluSerArgLeuArgGluSerLysValPheProSerGluIleThrAsp 1034
 Db 3781 AGTCTGTGCGAAGAGCCGATTTAGAGAAATCAAAAGTTTCCAGTGAATTAACAGAT 3840
 QY 1035 ThrValAlaAlaSerThrAlaGlnSerProGlyMetAsnLeuSerHisSerAlaSerSer 1054
 Db 3841 ACAGTTGCTGCTCTTACAGCTCAGAGCCCTGAAATGAATCTGTCTCACTGTCATCATCC 3900
 QY 1055 LeuSerLeuGlnGlnAlaPheSerGluLeuArgArgAlaGlnMetThrGluGlyProAsn 1074
 Db 3901 CTTAGTCTACACAGGCTTTTCTGAACTTGAAGGTGCGCAAAATGACAGAAAGACCCAA 3960
 QY 1075 ThrAlaProProAsnPheSerHisThrGlyProThrPheProValAlaProProPheLeu 1094
 Db 3961 ACAGACCTCCAACTTATGCTATACAGAGAACCAATTTCCAGTAGTACCTCTTCTTA 4020
 QY 1095 SerSerIleAlaGlyValProThrThrAlaAlaAlaThrAlaProValProAlaThrSer 1114
 Db 4021 AGTAGCAATTGCTGAGAGTCCCAACCAAGCAGCAGCAGCAGCAGCAGCTCTGCAACAGC 4080
 QY 1115 SerProProAsnAspIleSerThrSerValIleGlnSerGluValThrValProThrGlu 1134
 Db 4081 AGCCCTCTTAATGACATTTTCCACATCAAGTATTAAGTGTGAGTTACAGTCCCACTGAA 4140
 QY 1135 GluGlyIleAlaGlyValAlaThrSerThrGlyValValThrSerGlyLysLeuProIle 1154
 Db 4141 GAGGGAGATTGCTGAGTTGCCACACAGCAGCTGTGGTAACTTCAAGTGTCTCCCACTA 4200
 QY 1155 ProProValSerGluSerProValLeuSerSerValAlaSerSerIleThrIleProAla 1174
 Db 4201 CCACCTGTGCTGATACCAAGTACTTCCAGCTTCAAGTTTCAAGTATCACAAATCTGCA 4260
 QY 1175 ValValSerIleSerThrThrThrSerProSerLeuGlnValProThrSerThrSerGluIle 1194
 Db 4261 GTTGTCTCAATATTTACTACATCCCGCTCACTTCAAGTCCCAATCCCAATCTGAGATC 4320
 QY 1195 ValValSerSerThrAlaLeuLysProSerValThrValSerAlaThrSerAlaSerAla 1214
 Db 4321 GTTGTCTTCTAGTACAGACCTGATCTTCAAGTAAAGTTTCAAGTCCCAATCTGAGATC 4380
 QY 1215 GlyGlySerThrAlaThrProGlyProLysProProAlaValValSerGlnAlaAla 1234
 Db 4381 GGGGGCAGTACTGCTACCCAGAGTCTTCAAGCTTCAAGCTTCAAGTATCTCAGAGGAGCA 4440
 QY 1235 GlySerThrThrValGlyAlaThrLeuThrSerValSerThrThrThrSerPheProSer 1254
 Db 4441 GGGAGCACTACTGTGGAGGCCAATTAACTCAAGTTTCTCAACCACTTCAATCCCAAGC 4500
 QY 1255 ThrAlaSerGlnLeuSerIleGlnLeuSerSerSerThrSerThrProThrLeuAlaGlu 1274
 Db 4501 ACAGCTTCAACAGCTGTCACTTCAAGTTCAGAGAGTACTTCTACTCTTAACTTAAAGTAA 4560
 QY 1275 ThrValValValSerAlaHisSerLeuAspLysThrSerHisSerSerThrThrGlyLeu 1294
 Db 4561 ACCGTGTACTTGGCCAGCACTCACTAGTAAACATCTCATAGCAGTACCAACTGAGATTG 4620
 QY 1295 AlAPheSerLeuSerAlaProSerSerSerSerSerProGlyAlaGlyValSerSerTyr 1314
 Db 4621 GCTTCTCCTCTCTGCAACATCTTCTTCTCTCTCTCTCTGAGAGAGAGTCTAGTTAT 4680
 QY 1315 IleSerGlnProGlyLysLeuHisProLeuValIleProSerValIleAlaSerThrPro 1334
 Db 4681 ATTTCTCAGCTGTGGGCTGCACTCTTGTGATTCATCAAGATAGCTTCTACTCT 4740
 QY 1335 IleLeuProGlnAlaAlaGlyProThrThrThrProLeuLeuProGlnValProSerIle 1354
 Db 4741 ATCTTCTCCCAACAGAGAGACTTACTTCAACCTTATTAATCCCAAGTACTAGTATC 4800

QY 1355 ProProLeuValGlnProValAlaAsnValProAlaValGlnGlnThrLeuLeuHisSer 1374
DB 4801 CCACCTTGAGACGCTGTTGCAATGCTGCTGACAGACCACTAATTCATAGT 4860
QY 1375 GlnProGlnProAlaLeuLeuProAsnGlnProHisThrHisCysProGluValAspSer 1394
DB 4861 CAGCTCAACAGCTTGTCTCCCAACCAAGCCCACTCATTTGCTGAGTAGATTC 4920
QY 1395 AsnThrGlnProValAlaProGlyValLeuAspIleLeuThrLeuGlnGlnLeuLeuArg 1414
DB 4921 GATACACAAACCAAGCTCTGGAATTGATGACATTAAGACTCTAGAAGAAAGCTGGG 4980
QY 1415 SerLeuPheSerGlnHisSerSerSerGlyValGlnHisAlaSerValSerLeuGlnThr 1434
DB 4881 TCTCTGTTCACTGAACACAGCTCATCTGAGACTCAGCATGCTCTGCTCTCACTGGAAGCC 5040
QY 1435 SerLeuValIleGlnSerThrValThrProGlyValLeuProThrAlaAlaProSer 1454
DB 5041 TCACCTAGCATAGAGAGACAGTGTCAACACAGCATCCCACTACTGCTGTGACCAAGC 5100
QY 1455 LysLeuLeuThrSerThrThrThrThrCysLeuProProThrAsnLeuProLeuGlyThr 1474
DB 5101 AAACCTCCGACTTCTACCAACAGATCTTGCTTACCAACCACTTAACCACTAGAACCA 5160
QY 1475 ValAlaLeuProValThrProValValThrProGlyGlnValSerThrProValSerThr 1494
DB 5161 GTTGCTTTGCCAGTTACCAACAGTGTCACTGAGCAAGTTTCAACCCAGTCAAGCACT 5220
QY 1495 ThrThrSerGlyValLysProGlyThrAlaProSerLysProProLeuThrLysAlaPro 1514
DB 5221 ACTACATCAGAGGAGTAAACCTGGAACCTGCTCCCTCAAGCAACCTCTAATTAAGGCTCG 5280
QY 1515 ValLeuProValGlyThrGlnLeuProAlaGlyThrLeuProSerGlnGlnLeuProPro 1534
DB 5281 GTGCTGCAGAGGGAGTAACTTCCACAGAGTACTTCAACCAAGGACAGTGTGCACTT 5340
QY 1535 PheProGlyProSerSerLeuThrGlnSerGlnGlnProLeuGlnAspLeuAspAlaGlnLeu 1554
DB 5341 TTTCCAGAACCTTCTCTAACCCAGTCCAGCACTCTAGAGATCTTGATGCTCAATTG 5400
QY 1555 ArgArgThrLeuSerProGluValIleThrValThrSerAlaValGlyProValSerMet 1574
DB 5401 AGAAGAACACTTATGCTCAGAGATKATCAAGTGACTTGCAGGTGTGCTGTGCTCATG 5460
QY 1575 AlaAlaProThrAlaIleThrGlnAlaGlyThrGlnProGlnLysGlyValSerGlnVal 1594
DB 5461 GCGGCTCAACAGCATACAGAAACAGAAACAGCAAGCTCTAGAAAGGCTGTTCTCAAGTC 5520
QY 1595 LysGlnGlyProValLeuAlaThrSerSerGlyAlaGlyValPheLysMetGlyArgPhe 1614
DB 5521 AAAGAGAGCCCTGCTCTAGCACTAGTTCAGAGAGCTGTGTTTAAAGATGGAGCATTT 5580
QY 1615 GlnValSerValAlaAlaAspGlyValGlnLysGlnGlyLysAsnLysSerGlnAspAla 1634
DB 5581 CAGGTTTCTGTTGAGCAGACAGCTGCCAGAAAGAGGGTAAATAAGTCAGAAATATCA 5640
QY 1635 LysSerValHisPheGlnSerSerThrSerGlnSerValLeuSerSerSerSerPro 1654
DB 5641 AAGTGTGTCATTGTAATCCAGCACTCAGAGTCTCAGTCTCATCAGTATGAGTCA 5700
QY 1655 GlnSerThrLeuValLysProGlnProAsnGlyIleThrIleProGlyIleSerSerAsp 1674
DB 5701 GAGATTAACCTTGTAACCAAGCCGAATGGCATTAACCATCCCTGTAATCTCTCAAGT 5760
QY 1675 ValProGlnSerAlaHisLysThrThrAlaSerGlnAlaLysSerAspThrGlyGlnPro 1694
DB 5761 GTGCAGAGAGTGCACAAACTACTGCTCAGAGGCAAAAGTCAAGACACTGGAGAGCT 5820
QY 1695 ThrLysValGlyArgPheGlnValThrThrThrAlaAsnLysValGlyArgPheSerVal 1714
DB 5821 ACCAAGGTTGACGTTTCAAGGTGACAACTACAGCAAAAGAGTGGGTGTTCTCTGTA 5880
QY 1715 SerLysThrGlnAspLysIleThrAspThrLysGlnGlyProValAlaSerProPro 1734

DB 5881 TCATAAACTAGAGCAAGATCTACGACACAAAGAAAGAGCAAGTGCATCTCTCTCT 5940
QY 1735 PheMetAspLeuGlnGlnAlaValLeuProAlaValIleProLysLysGlnLysProGlu 1754
DB 5941 TTTATGATTTGGAAACAGCTGTTCTTCTGCTGTGATACCAAAAGAAAGAAACCTGAA 6000
QY 1755 LeuSerGlnProSerHisLysAsnGlyProSerSerAspProGluAlaAlaPheLeuSer 1774
DB 6001 CTGTCAAGCTTCAACATCTTAATGGCCGTCTTCTGACCCGAGAGCCGCTTTTAAAGT 6060
QY 1775 ArgAspValAspAspGlySerGlySerProHisSerProHisGlnLeuSerLysSer 1794
DB 6061 AGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6120
QY 1795 LeuProSerGlnAsnLeuSerGlnSerLeuSerLysSerPheAsnSerSerLysMetSer 1814
DB 6121 CTTCCTAGCCAGAAATCTTAAGTCAAAAGCTTAATTAATTAATTAATTAATTAATTA 6180
QY 1815 SerAspAsnGlnSerAspIleGlnAspGlnAspLeuLysLeuGlnLeuArgArgLeuArg 1834
DB 6181 AGCCAAATAGTCAAGATATGAAAGATGAAAGATTAAGTTAGAGCTGCAAGCTAGCA 6240
QY 1835 AspLysHisLeuLysGlnIleGlnAspLeuGlnSerArgGlnLysHisGlnIleGlnSer 1854
DB 6241 GATAAATCTCAAAAGAGATTCAGGACCTGCAAGTGCAGAGTGCAGAAATTAATTAAT 6300
QY 1855 LeuThrThrLysLeuGlyLysValProProAlaValIleLeuProProAlaAlaProLeu 1874
DB 6301 TTTATATCCAAACCTGGGCAAGGTCCTCCCTGCTGTTATATATATATATATATATAT 6360
QY 1875 SerGlyArgArgArgArgProThrLysSerLysGlySerLysSerSerArgSerSerSer 1894
DB 6361 TCAGGAGAACACACACACACACACACACACACACACACACACACACACACACACAC 6420
QY 1895 LeuGlyAsnLysSerProGlnLeuSerGlyAsnLeuSerGlyGlnSerAlaAlaSerVal 1914
DB 6421 TTGGGAAATAAAGCCCTCCAGCTTCAAGTAACTGCTGTGCTCAGAGTGCAGCTTCACTC 6480
QY 1915 LeuHisProGlnGlnThrLeuHisProProGlyAsnIleProGlnLysGlnGlnAsnGln 1934
DB 6481 TTGACCCCTCAGAGACCTCTCCCTCTGCAACATCCCAAGTCCGGGCAAGATAG 6540
QY 1935 LeuLeuGlnProLeuLysProSerProSerSerAspAsnLeuLysSerAlaPheThrSer 1954
DB 6541 CTGTTACAGCCCTTAAGCCATCTCCCTCAGTGAACACCTTAATTCAGCTTCAAGT 6600
QY 1955 AspGlyAlaIleSerValProSerLeuSerAlaProGlyGlnGlyThrSerSerThrAsn 1974
DB 6601 GATGGTCCATTTCACTACGACCAAGCTTCTGCTCAGAGTCAAGAACACAGACGCAAAAC 6660
QY 1975 ThrValGlyAlaThrValAsnSerGlnAlaAlaGlnIleGlnProProAlaMetThrSer 1994
DB 6661 ACTGTTGGGGCAACGTAACAGCCCAAGCCCAAGCTCAGCTCTGCAAGTCAAGTCC 6720
QY 1995 SerArgLysGlyThrPheThrAspAspLeuHisLysLeuValAspAsnTrpAlaArgAsp 2014
DB 6721 AGCAGGAAGGCACTTCAAGATGACTTGCACAAAGTTGTAAGACAAATGGGCCGAGAT 6780
QY 2015 AlaMetAsnLeuSerGlyArgArgLysSerLysGlyHisMetAsnTrpGlnGlyProGly 2034
DB 6781 GCCATGAATCTCTCAGGAGGAGAGAGAAAGCAAGGCACTGAATTAAGAGGCCCCCTGA 6840
QY 2035 MetAlaArgLysPheSerAlaProGlyGlnLeuCysIleSerMetThrSerAsnLeuGly 2054
DB 6841 ATGGCAAGAAAGTTCTTGCACCTGGGCACTGTCACTTCATCATCACTCAAGCTGGAT 6900
QY 2055 GlySerAlaProIleSerAlaAlaSerAlaThrSerLeuGlyHisPheThrLysSerMet 2074
DB 6901 GGCTCTGCCCATCTCTGCAAGCACTAGTACTCTTCAAGTCACTTCAACCAAGTCTAG 6960
QY 2075 CysProProGlnGlnThrGlyPheProAlaThrProPheGlyAlaGlnTrpSerGlyThr 2094

Db 6961 TGCCCCCAGACAGATGATGGCTTCCAGCTACCCCATTTGGCGCTCATGATGAGTGAGACG 7020

Qy 2055 GYGLYPRGIALPProGlnProLeuGlyGlnPheGlnProValGlyThrAlaSerLeuGln 2114

Db 7021 GGTGGCCCGACGACACAGACCACTTGCGCGATTCCAACCTGGGAAGTCCCTCTTGACG 7080

Qy 2115 AapPheanPILSerAsnLeuGlnLysSerIleSerAsnProProGlySerAsnLeuGly 2134

Db 7081 AATTTCACATCAGCAATTTGAGAAATTCATCAGCAACCCCGAGCTCCCACTGCGG 7140

Qy 2135 ThrThr 2136

Db 7141 ACCACT 7146

RESULT 5

LOCUS HSA296290 7149 bp mRNA linear PRI 17-MAR-2003

DEFINITION Homo sapiens mRNA for putative protein kinase (WNK1 gene).

ACCESSION AJ296290

VERSION AJ296290.1 GI:1125347

KEYWORDS protein kinase; WNK1 gene.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

AUTHORS Verissimo, F. and Jordan, P.

TITLE WNK kinases, a novel protein kinase subfamily in multi-cellular organisms

JOURNAL Oncogene 20 (39), 5562-5569 (2001)

MEDLINE 21455683

PUBMED 11571656

REFERENCE 2 (bases 1 to 7149)

AUTHORS Jordan, P.

TITLE Direct Submission

JOURNAL Submitted (02-NOV-2000) Jordan P., Centro de Genetica Humana, Instituto Nacional de Saude 'Dr. Ricardo Jorge', Avenida Padre Cruz, Lisboa, 1649-016, PORTUGAL

FEATURES

source

1..7149

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/chromosome="12"

/map="12p13.3"

/tissue_type="heart"

1..7149

/gene="WNK1"

1..7149

/gene="WNK1"

/codon_start=1

/evidence="experimental"

/product="putative protein kinase"

/protein_id="CAC15059.1"

/db_xref="GI:1125348"

/db_xref="GOA:Q9H4A3"

/db_xref="SPTREMBL:Q9H4A3"

/translation="MSGGAAEKOSTSTPSGLFLSPAPAPKXSSDSSVGEKLGAAA
DAYVGRTEERYRRRRHTMDXSRGAAATTTTHERFFPRRSVICDSNATLRLPLST
POPSIPAAYPOSAAPREPHREETATATSOVAOOPRAAAGPBAVAPASTVPST
SKORPVQPSLVGSKKEPPRPARSGSGGSAKEPDEERSQDDDI EELETAVGMSNG
RFLKFDLEIGRSFTVYKGLDTEETVEVAMCELODKRLTKSERQKPEAEMLKGLQ
HPNIVRFYDSEWSTVKGKCIIVATLMTSGTLTKYLPKVMKIKVLRAEMCQILKG
LOPLHRTPTPIHRLDKCNI FITGPTGSVKIGLGLATLRASFAKSVIGTEPMAP
EMEEKDESVDVYAFGCMLEMASTSEPYSECONAOIYRATSGVYKPSAPDKVAP
EVEETIEGCIKONKDERYSTKDLINHA.FPOERTGVRVELAEEDDEKIAIKLRLED
IKLKGYKKNENAIERSFDLERDVEDVAGEMESGIVCEGDKHTAKAIKDVSLIK
RKERQRLVREBGRKKOESSLKOVEGSSAQGTGKLPASGTPIASTTASVS
TOVEPEPEADQHOLOQOOPISIVLSDGVDSGQSSVFTESRVSSQOTVSYSQSH
QASHTVPHGISVTOAOSQPHGVVPSVVAOSQSGSSSLTGVSSQPIOHQ
QOOGIOCTAPPOOTVYSLSTSTSSAATTAOPVPOAPOVLPVYASGKOLVSOV
PTIOGEFOIVATOPSVVPHSGAHPLPGVGLPTPLLPYVPSQIPISPHSTANT
GFSRLPTMAAGITPLTLTASSATITATIGVSTVPSQPLTLQLPVTQLPSOVHQL

LOPAVOSMGIPANLGOAEVLESLSSGDVLYOGFPPLPPOYPGDSNIAPSSNVAVCIH
STVLSPEPTEVLATPGYFPTVVOVYVSNLLVPMGVGVOVONSOPGASIAQAPTTSS
SQOVLSTOGVSOVAPAPVAVAPQAPQOTTLASSVDSASHDVASMGDSNENVPSS
SSGHBERTTGRHTRKSVRSKSRHEKTSRPLKRLILANSKNGDRIVECOLLETNRKMT
FKPLDGDNPEELATIMVNDPILATIERESSVDQREIIEKADEMLSDVSEVPEGDO
GLESLIOGKDDEYFGFSQKLEGEFKOPIPASMPQOIGIPTSLSLTQVSHVAGVRFLVSP
VPSRLRESKVPSEIIDTVNASTAOSGKMLSHSASLSLOAFSETRRAAMGTGPN
TAPNFSHTGTFPVVPEPLSIAGVPTTAATAPVPATSSPNDISNVSIOEYTV
TEEGIAGVATSTGYVTGGALPIPVVSEPVLSVVSSTITTAIVASISITTSBLQVPTS
TSBIVSSTALPESVTSATSASAGSGTATGPPPAVNSQOAGSTVGNATLSVST
TTSFSTASQSLQSSSTSTPTLAETVVASHLDKTSHSTGLASLAPSSSS
PGAQVSSVYISQPGHPLVIPSVIASPTLPOAGAPSTPLIPVPSIPVQVPAVY
PAVOOTLIHSQOPALPNOPHTHCPEVSDPTOKAPCIDIDIKTLEKLRSEFSSHS
SGAQSASLETSLVISSTVPGIPTTAIVAPSKULITSTSTCLPTNPLCTVALLPVT
PVYTRGQVSTSTTSGVRCGTPAASKRPLKAVLVGTELPAGTLPSEDLPPPPG
SLTOSQPLELDQOLRTTSSPEIITVTSAGVSMAPVATITAGTQPOGVQVKE
GPVATISSGAGVFMGRFOVSVADGQKEKNSSEDAKSHFESSSTESSVLSSSP
ESTLVKEPNEIIPGISDVPESAHKTASSEAKSDGQPKVGRFOVTYANKGRF
SVSKTEBKTIDTKKEGVASPPMDLOAVLPAVIPKERELESEPSHLNPSDPEA
AFLSRDVGDSGSPHSHOLSKSLPSONLSOSLSNSFNSYMSDNDSODEDDIKL
ELRLRDKHLKEIODLOSROKHEIESTYTLGKVPVAVITIPPAPLSGRRRRPTKSKG
SKSRSSSLGKKSQLSGNLSGGSASVLEHQULHPPGNIPESGQMLDLPKPSBS
SDNLYSAFTSDGALSVPSLAPGGGTSTNTVATVNSQAAQAPPMATSSRKGTFTD
DLHLVDMNARADANLISGRGSKAMNYEGMARKEAPGOLCISMTSNGSGAPIS
AASATISLGHFTKSMCPQOYGFPATPFGAQSSTGCGAPQLGQFQFVGTASLQENFI
SNLQKISNPGSNLRTT"

1..759

/gene="WNK1"

/number=1

760..932

/gene="WNK1"

/number=2

933..1153

/gene="WNK1"

/number=3

1154..1311

/gene="WNK1"

/number=4

1312..1400

/gene="WNK1"

/number=5

1401..1620

/gene="WNK1"

/number=6

1621..1951

/gene="WNK1"

/number=7

1952..2139

/gene="WNK1"

/number=8

2140..2223

/gene="WNK1"

/number=9

2224..2373

/gene="WNK1"

/number=10

2374..2832

/gene="WNK1"

/number=11

2833..3111

/gene="WNK1"

/number=12

3112..3309

/gene="WNK1"

/number=13

3310..3372

/gene="WNK1"

/number=14

3373..3489

/gene="WNK1"

/number=15

3490..3663

/gene="WNK1"

exon

exon /number=16
 3664. .3721
 /gene="WNK1"
 /number=17
 3722. .3844
 /gene="WNK1"
 /number=18
 3845. .5280
 /gene="WNK1"
 /number=19
 5281. .5364
 /gene="WNK1"
 /number=20
 5365. .5448
 /gene="WNK1"
 /number=21
 5449. .5509
 /gene="WNK1"
 /number=22
 5510. .5583
 /gene="WNK1"
 /number=23
 5584. .6245
 /gene="WNK1"
 /number=24
 6246. .6448
 /gene="WNK1"
 /number=25
 6449. .6643
 /gene="WNK1"
 /number=26
 6644. .6831
 /gene="WNK1"
 /number=27
 6832. .7149
 /gene="WNK1"
 /number=28

ORIGIN

Alignment Scores:

Pred. No.: 1.17e-254 Length: 7149
 Score: 10677.00 Matches: 2135
 Percent Similarity: 89.63% Conservative: 1
 Best Local Similarity: 89.63% Mismatches: 246
 Query Match: 98.75% Indels: 1
 Gaps: 1

US-10-010-720-14 (1-2136) X HSA296290 (1-7149)

QY 1 MetSerGIyGIyAlaIaIaGIyGlnSerThrProGIySerLeuPheIeuSerPro 20
 Db 1 ATGTCTGGCGGGCGGCAAGAGAGACACACTCCGGTTCCTGTTCTCTCGCG 60
 QY 21 ProAlaProAlaProLyAsnGIySerSerSerSerSerValGIyGluIySLeuGIy 40
 Db 61 CCGGCTCTGCCCCCAAGATGGCTCCAGCTCCGATTCCTCGTGGGGAGAACTGGGA 120
 QY 41 AlaAlaAlaAlaAspAlaValThrGIyArgThrGIyGluIyTyrArgArgArgHisThr 60
 Db 121 GCCGGCGCGCGCGACGCTGTGACCGCGACGACCGAGGTACAGGCGCGCGCCACACT 180
 QY 61 MetAspLyAspSerSerArgIyAlaIaIaIaIaThrThrThrThrThrGluHisArgPhePhe 80
 Db 181 ATGGACAAGAGACGCGTGGGGCGCGCGACCACTACCACTGAGCACTGGCTTCTTC 240
 QY 81 ArgArgSerValIleCysAspSerAsnAlaThrAlaLeuGluLeuProGIyLeuProLeu 100
 Db 241 CCGCGGAGGCTATCTGGACCTCCAAATGCCACTGGCGTGGAGCTTCCGGCTTCTCTT 300
 QY 101 SerLeuProGIyProSerIleProAlaAlaValProGlnSerAlaProProGIyProHis 120
 Db 301 TCCCTGCCCAAGCCAGCATCCCGCGGCTGTCCGACAGTGTCTCAACCGAGGCCCGCAC 360

QY 121 ArgGIyGluThrValThrAlaThrAlaThrSerGlnValAlaGlnGlnProProAlaAla 140
 Db 361 CGGAGAGAGACGTTGACCGGCCACCGCACTTCCAGGTAGCCGACGCGCTCGAGCGCT 420
 QY 141 AlaAlaProGIyGlnGlnAlaValAlaGIyProAlaProSerThrValProSerSerThr 160
 Db 421 GCCGCCCTGGGGAACAGCCCGCTCGCGGCTCGCCCTCGACTGTGCCAGAGTACC 480
 QY 161 SerLyAspArgProValSerGlnProSerLeuValGlySerIySLeuGIyProProPro 180
 Db 481 AGCAAAAGCCGCCAGTGTCCAGCTTGTGGGGACAAAGAGAGAGCCGCGCG 540
 QY 181 AlaArgSerGIySerGIyGIySerAlaIySLeuProGlnGlnGluArgSerGlnGln 200
 Db 541 GCGAAGATGGACGCGCGCGCGGACCGCAAGAGCCACAGAGAGAGAGAGCCAGCAG 600
 QY 201 GlnAspAspIleGIyGluLeuGluThrIySAlaValGIyMetSerAsnAspGIyArgPhe 220
 Db 601 CAGATGATATCGAAGAGCTGAGACCAAGCGCTGGGGAATGTCTAACATGAGCGCTTT 660
 QY 221 LeuIySPhaAspIleGIyIleGIyArgGIySerPheIySThrValIyIySLeuAsp 240
 Db 661 CTCAAGTTTGCATCGAAATCGCAAGGCTCTTTAGACGCTACAAAGTGTGAC 720
 QY 241 ThrGIyThrThrValGIyValAlaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIa 260
 Db 721 ACTGAACCAACCGTGAAGTGGCTGTGTGAATCGACGAGATCGAAATTAACAAGTCT 780
 QY 261 GluArgGlnArgPheIySLeuGluAlaGluMetLeuIySLeuGlnHisProAsnIle 280
 Db 781 GAGAGCGCAGATTTAAAGAAAGAGCTGAATGTAAAGGTCTTCACATCCCAATTT 840
 QY 281 ValArgPheIyAspSerSerThrGlnSerThrValIySLeuIySLeuValIleValIle 300
 Db 841 GTTAGATTATATGATCTCTGGGAATCCACAGTAAGAAAGAAAGATGATCTTTGGTG 900
 QY 301 ThrGluLeuMetThrSerGIyThrLeuIySThrIyLeuIySArgPheIyValMetIyS 320
 Db 901 ACTGAAGTTATGAGCTGTGGAACACTTAACGATATCGAAAGGTTTAAAGTATGATGAG 960
 QY 321 IleIySValLeuAspSerSerThrCysArgGlnIleLeuIySLeuGlnPheLeuHisThr 340
 Db 961 ATCAAAAGTTCAAGAAAGCTGTGCGGTGCGAGTCTTAAGGTCTTCAATTTCTCAATCT 1020
 QY 341 ArgThrProProIleIleHisArgAspLeuIySLeuIySLeuIySLeuIySLeuIySLeu 360
 Db 1021 CGAAGTCACTATCATTCACCGCGCATCTTAAGTGAACATCTTATACCGGCGCT 1080
 QY 361 ThrGlySerValIySLeuIySLeuIySLeuIySLeuIySLeuIySLeuIySLeuIySLeu 380
 Db 1081 ACTGGCTAGTCAAGATTTGAGACCTCGGTCTGGCAACCTGAAACCGGCTTCTTGGCC 1140
 QY 381 LySerSerValIleGIyThrProGluPheMetAlaProGluMetIyGluIySLeuIySLeu 400
 Db 1141 AAGAGTGTGATGATGACCCCAAGATTCATGGCCCTGAGATGATAGAGAGAAATATGAT 1200
 QY 401 GluSerValAspValTyrAlaPheGIyMetCysMetLeuGluMetAlaThrSerGIyIy 420
 Db 1201 GATTCGTTGAGCTTATGCTCTTGGAGTGTGACATGCTTGAATGGCTACATCTGAATAT 1260
 QY 421 ProTyrSerGIyCysGlnAsnAlaAlaGlnIleTyrArgValIyThrSerGIyValIyS 440
 Db 1261 CTTTACTGGAGTGCCAAAATCTCGAGATCTACCTCGGCTGACCAAGTGGGTGAAG 1320
 QY 441 ProAlaSerPheAspLyValAlaIleProGluValIySLeuIleIleGluIyCysIle 460
 Db 1321 CCAGCAGATTTGACAAAGATGCAATCTCTGAAGAGAGAAATTAATGAAGATGACATA 1380
 QY 461 ArgGlnAsnIySAspGIyArgTyrSerIleIySAspLeuLeuAsnHisAlaPhePheGln 480
 Db 1381 GCACAAACCAAGAGTAAGATATTCATCAAAAGCTTTTGAACCAAGCTTCTTCCAA 1440
 QY 481 GluGluThrGluValArgValGluLeuAlaGluIyAspAspGIyGluIySLeuIleAlaIle 500

```

Db      1441 GAGGAAACAGAGATCGGCTAGCAATTTAGACAGAAAGATGATGAGAAAAAATACCCATA 1500
QY      501  LysLeuTPrLeuArgIleGluAspIleLysIleLysGlyLysTyrLysAspAsnGlu 520
Db      1501 AAATTATGGCTACGATTGAAATTTAAGAAATTTAAGGAAAAATACAAAGATATATGAA 1560
QY      521  AlaIleGluPheSerPheAspLeuGluArgAspValProGluAspValAlaGlnIleMet 540
Db      1561 GCTATGTAGATTCTTTTGTATTAGAGAGAGATGTCACAGAAAGATGTTGACAAAGAAATG 1620
QY      541  ValGluSerGlyTyrValCysGlnGluAspHisLysThrMetAlaLysAlaIleLysAsp 560
Db      1621 GTAGAGTGGGTATGTCGTGTGAAGGTATCAACAGACCATGAGCTTAAGGCTATCAAAAC 1680
QY      561  ArgValSerLeuIleLysArgLysArgGluGlnArgGlnLeuValArgGlnGlnGln 580
Db      1681 AGAGATCATTAATTAAGAGAAAGAGAAAGAGACGCGCATGTTGTTACGGAGAGAGCAAGAA 1740
QY      581  LysLysLysGlnGluGluSerSerLeuLysGlnGlnValGlnGlnSerSerAlaSerGln 600
Db      1741 AAAAAAAGCAGAGAGAGAGAGAGAGCTCAACAGCAGGTAGAAACAATCCAGTGCCTCCAG 1800
QY      601  ThrGlyIleLysGlnLeuProSerAlaSerThrGlyIleProThrAlaSerThrThSer 620
Db      1801 ACAGGAATCAAGCAAGCTCCCTTCTGTAGACACCGGCATACCTACTGCTTTCACCACTTCA 1860
QY      621  AlaSerValSerThrGlnValGluProGluGluProGluAlaAspGlnHisGlnGlnLeu 640
Db      1861 GCTTCAGTTTCTACACAGATGAGAACCTGACAGACCTGAGGACAGATCAACATCAACACTA 1920
QY      641  GlnTyrGlnGlnProSerIleSerValLeuSerAspGlyThrValAspSerGlyGlnGly 660
Db      1921 CAGATACCAAGCAACCAAGATATATCTGTATTCTGATGGGACGTTGACAGTGGTCAAGGGA 1980
QY      661  SerSerValPheThrGluSerArgValSerSerGlnGlnThrValSerTyrGlySerGln 680
Db      1981 TCCTCTGCTTTCACAGAAATCTCGAGTGAAGCAGCCACAGACAGACAGTTTCAATATGTTCCAA 2040
QY      681  HisGlnGlnAlaHisSerThrGlyThrValProGlyHisIleProSerThrValGlnAla 700
Db      2041 CATGAACAGGCACATTTCTACAGGACAGCATGCCAGGACATATCTTCTCTGTCACAGCA 2100
QY      701  GlnSerGlnProHisGlyValTyrProProSerSerValAlaGlnGlyGlnSerGlnGly 720
Db      2101 CAGCTCAGGCCCATGGGGTATATCCACCTCAAGTGGGACAGGGGACAGGCCAGGCT 2160
QY      721  GlnProSerSerSerSerLeuThrGlyValSerSerSerGlnProIleGlnHisProGln 740
Db      2161 CAGCCATCTCTAGATAGCTTAAACAGGGGTTTCATCTTCCCAACCCATACAAACATCCCTCAG 2220
QY      741  GlnGlnGlnGlnGlyIleGlnGlnThrAlaProProGlnGlnThrValGlnTyrSerLeuSer 760
Db      2221 CAGCAGCAGGGAATACAGCAGACAGACGCCCTCTCTCAACAGACAGTACAGTATTCACCTTCA 2280
QY      761  GlnThrSerThrSerSerGluAlaThrThrAlaGlnProValSerGlnProGlnAlaPro 780
Db      2281 CAGACATCAACCTCCAGTGAAGGCCACTACTGACAGCAGCAGTGAAGTCAAGCTCAAGCTCCA 2340
QY      781  GlnValLeuProGlnValSerAlaGlyLysGln----- 791
Db      2341 CAAGTCTTGCTCAAGATACAGCTGGGAAAACAGCTTCCAGTTCCAGCAGATACCAACT 2400
QY      791 ----- 791
Db      2401 ATCCAAAGGCGAAGCTCAGATCCAGTTGGCAGACACAACCTCGGTTGTTCCAGTCCACTCT 2460
QY      791 ----- 791
Db      2461 GGTCGCTATTCTTCAGTGGAGACAGCCGCTCCTACTCCTTGCTCCTCAGTACCTT 2520
QY      791 ----- 791

Db      2521 GTCTCTCAGATTCCCATATCAACTCTCATGTGTCTACGGGCTCAGACAGGTTTCTCATCC 2580
QY      791 ----- 791
Db      2581 CTTCCTCATCAATAGCAGCTGGCATTAAGCTCAGCTCTGTCTCAGCTGGCTTCAATCTGCT 2640
QY      791 ----- 791
Db      2641 ACAAAGCTGGATCCCGGGGGTATCAACTGTGTTCTTCTAGTCAAGCTTCCAAACCTTTTG 2700
QY      791 ----- 791
Db      2701 CAGCCTGACTCAGCTGCAAGTCAGGTTCAACCAAGCTCTCTCAACAGCAAGCACTTGAG 2760
QY      791 ----- 791
Db      2761 TCATGGGAATACAGACTAACCTTGGACAAAGCTGTGAGTTCCACTTCTCTGGAGAT 2820
QY      791 ----- 791
Db      2821 GTTCTGTACAGGGCTTCCCACTTGACTGCAACCAAGTACCCAGAGATTCAAATATT 2880
QY      791 ----- 791
Db      2881 GCTCCCTTCTTCAAGTGGCTTCTGTGTGATTCATTCATCTACAGTCCATCCCTCCAGT 2940
QY      791 ----- 791
Db      2941 CCGACAGAAATACAGCTGCTACACCTGGGTACTTCCACAGTGTGAGCAGCTTATGTGAA 3000
QY      791 ----- 791
Db      3001 TCAATCTTTTATGTTCTTATGGGTGTGTAGAGAGACAGTTCAAGTGTCCAGCCACGGA 3060
QY      792 ----- SerThrGln 794
Db      3061 GGGAGTTTAGACAGACGCCCACTACCTCCACAGACAGATTTTGGAGATATCTCAG 3120
QY      795  GlnValSerGlnValAlaProAlaGluProValAlaValAlaGlnProGlnAlaThrGln 814
Db      3121 GAGGTCTCTAGGTGCTCTCTGAGAGCAGTTGACATGACACAGCCCAAGCTACACAG 3180
QY      815  ProThrThrLeuAlaSerSerValAspSerAlaHisSerAspValAlaSerGlyMetSer 834
Db      3181 CCGACCACTTTGGCTTCTCTGTAGACAGTGCACATTCATGTTGCTTCAAGATATAGT 3240
QY      835  AspGlyAsnGluAsnValProSerSerSerGlyArgHisGlnGlyArgThrThrLysArg 854
Db      3241 GATGGCAATGAGAAAGTCCCATCTTCCAGTGAAGGATGAAGGAAGAACCTACAAAACGG 3300
QY      855  HisTyrArgLysSerValArgSerArgSerArgHisGlnLysThrSerArgProLysLeu 874
Db      3301 CATTTACGAAAATCTGTAAAGAGATCGCTCTCGACATGAAAAAACTTCACGCCCAAAATTA 3360
QY      875  ArgIleLeuAsnValSerAsnLysGlyAspArgValValGluCysGlnLeuGluThrHis 894
Db      3361 AGAATTTTGAATCTTTCATTAAGAGACCGAGTGTAGAAATGCTCAATTAAGACCTCAT 3420
QY      895  AsnArgLysMetValThrPheLysPheAspLeuAspGlyAspAsnProGlnIleAla 914
Db      3421 AATAGAAAAATGGTTACATTCAAATTTGACTGATGTGTGACAAACCCGAGGAGATAGCA 3480
QY      915  ThrIleMetValAsnAsnAspPheIleLeuAlaIleGluArgGluSerPheValAspGln 934
Db      3481 ACAATTTATGTTGAACATGACTTATTTAGCAATATGAGAGAGAGAGTGTGTGATCA 3540
QY      935  ValArgGluIleIleGluLysAlaAspGluMetLeuSerGluAspValSerValGluPro 954
Db      3541 GTGCGAGAAATTAATGAAAAAGCTGATGAATTCCTAGTGAAGATCTCAGTGTGAAACCA 3600
QY      955  GluGlyAspGlnGlyLeuGluSerLeuGlnGlyLysAspAspTyrGlyPheSerGlySer 974
Db      3601 GAGGGTATCAGGATTTGAGAGTCTACAAAGAAAGATGACTATGCTTTTCAAGGTTCT 3660

```

QY	975	GlulysLeuGluGluGluPheLeuProIleProAlaSerSerMetProGlnGlnIle	994
DB	3661	CAGAAATTTGAGAGAGGTTCAACCAACAAATTCCTGCTTCACAGCCACAGCAATA	3720
QY	995	GlyIleProThrSerSerLeuThrGlnValAlaHisSerAlaGlyArgAlaPheIleVal	1014
DB	3721	GGCATTCCTACCACTTCTTACCAAGTTGTTCAATTCGCGGGAAGCGGTTATAGTG	3780
QY	1015	SerProValProGluSerArgLeuArgGluSerIleValPheProSerGluIleThrAsp	1034
DB	3781	AGTCTGTGGCAGAAAGCCGATTCAGAAATCAAAAGTTTCCCAAGTAAATACAGAT	3840
QY	1035	ThrValAlaAlaSerThrAlaGlnSerProGlyMetLeuLeuSerHisSerAlaSerSer	1054
DB	3841	ACAGTTCTGCTCTTACAGCTCAGAGCCGGAATGAACTTGCTCACTGCACTCAATCC	3900
QY	1055	LeuSerLeuGlnGlnAlaPheSerGluLeuArgArgAlaGlnMetThrGluGlyProAsn	1074
DB	3901	CTTAGTCTACCAAGGCTTTTCTGAACTTAGAGTGGCCAAATGACAGAAAGCCCAAC	3960
QY	1075	ThrAlaProProAsnPheSerHisThrGlyProThrPheProValValProPheLeu	1094
DB	3961	ACAGACCTCCAACTTAGCATATACAGCAACCAATTCCAGTAGTACCTCTTCTTA	4020
QY	1095	SerSerIleAlaGlyValProThrThrAlaAlaThrAlaProValProAlaThrSer	1114
DB	4021	AGTAGCATGCTGAGAGTCCCAACCAAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC	4080
QY	1115	SerProProAsnAspIleSerThrSerValIleGlnSerGluValThrValProThrGlu	1134
DB	4081	AGCCCTCTTAATGACATTCCACATCACTAATTCAGTGTGAGGTAGAGTCCCACTGA	4140
QY	1135	GluGlyIleAlaGlyValAlaThrSerThrGlyValValThrSerGlyLeuProIle	1154
DB	4141	GAGGGGATGCTGAGGTGCCACAGCAGCAGGTGGTACTTCAAGGTGCTCCCACTA	4200
QY	1155	ProProValSerGluSerProValLeuSerSerValIleSerSerIleThrIleProAla	1174
DB	4201	CCACCTGTGCTGATCAACAGTACTTTCAGCCGATGATTCAGATACAAATCACTGCA	4260
QY	1175	ValValSerIleSerThrThrSerProSerLeuGlnValProThrSerThrSerGluIle	1194
DB	4261	GTTGCTCATATCTACTACATCCCCGTCATTCAGTCAAGTCCCACTCACTGAGATC	4320
QY	1195	ValValSerSerThrAlaLeuThrProSerValThrValSerAlaThrSerAlaSerAla	1214
DB	4321	GTTGTTTCTAGTACAGCACTGATCTTCAGTACAGTTCAGCAACTTCAGCTCTGCA	4380
QY	1215	GlyGlySerThrAlaThrProGlyProIleProAlaAlaValSerGlnGlnAlaIle	1234
DB	4381	GGGGGCAAGTACGTAACCCAGGTCTTAAGCTCCAGCTGATGATCTCAGCAGCAGCA	4440
QY	1235	GlySerThrThrValGlyAlaThrLeuThrSerValSerThrThrThrPheProSer	1254
DB	4441	GGCAGCATCTAGTGTGGAGCCCATTAACATAGTTTCCACACCACTTCATTCACAAG	4500
QY	1255	ThrAlaSerGlnLeuSerIleGlnLeuSerSerThrSerThrProThrLeuAlaGlu	1274
DB	4501	ACAGCTTCAGAGCTGCTCATTAGCTTACAGCAGTACTTCACTCACTTACTGAA	4560
QY	1275	ThrValValValSerAlaHisSerLeuAspIleThrSerHisSerSerThrThrGlyLeu	1294
DB	4561	ACCGGTAGTATGAGCAGCATCTCAGTAAAGATTCATAGCAGTACAGTGGATTG	4620
QY	1295	AlaPheSerLeuSerAlaProSerSerSerSerProGlyAlaGlyValSerSerThr	1314
DB	4621	GCTTCTCTCCCTCTGCAACATCTTCTCTCTCTCTGAGAGGAGATGTCTAGTAT	4680
QY	1315	IleSerGlnProGlyGlyLeuHisProLeuValIleProSerValIleAlaSerThrPro	1334
DB	4681	ATTCTCAGCTGGTGGCTGATCTTTGTCATTCATCAGTATAGTACTTCTACTCT	4740
QY	1335	IleLeuProGlnAlaAlaGlyProThrSerThrProLeuLeuProGlnValProSerIle	1354
DB	4741	ATTCTTCCCAAGCAGAGACTTACTTACACTTATATATCCCAAGTACCTAGATC	4800
QY	1355	ProProLeuValGlnProValAlaAsnValProAlaValGlnGlnThrLeuIleHisSer	1374
DB	4801	CCACCTTGATACAGCTGTGGTCCAAATGTGCTGCTTACAGCAGACATCAATTCATAGT	4860
QY	1375	GlnProGlnProAlaLeuLeuProAsnGlnProHisThrHisCysProGluValAspSer	1394
DB	4861	CAGCTTCACAGCTTGTGCTTCCCAACAGCCCAATCTCATTTGCTGAGTATGATTC	4920
QY	1395	AspThrGlnProIleValAlaProGlyIleAspAspIleThrThrLeuGlnGlyLeuArg	1414
DB	4921	GATACACACCCAAAGCTCTGGATTTGATGACATTAAGACTTAGAAGAAAGAACTGGG	4980
QY	1415	SerLeuPheSerGluHisSerSerSerGlyAlaGlnHisAlaSerValSerLeuGlnThr	1434
DB	4981	TCTCTGTTCAGTAAACAGCTCATCTGAGCTCAGATGCTGCTGCTCACTGAGACC	5040
QY	1435	SerLeuValIleGlnSerThrValThrProGlyIleProThrThrAlaValAlaProSer	1454
DB	5041	TCACTAGCTATAGAGAGCATGTCACACAGCATCCCACTACTGCTGTGACCAAGC	5100
QY	1455	LeuLeuLeuThrSerThrThrSerThrCysLeuProProThrAsnLeuProLeuGlyThr	1474
DB	5101	AAACTCTGACTTACCAACAGTACTTGTACCAACCAACATTAACCTAGAGAAC	5160
QY	1475	ValAlaLeuProValThrProValValThrProGlyGlnValSerThrProValSerThr	1494
DB	5161	GTTGCTTGGCAGTAAACAGTGTGACACCTGGGAGAAAGTTTACCCCACTGAGCACT	5220
QY	1495	ThrThrSerGlyValIleProGlyThrAlaProSerIleProLeuThrValAlaPro	1514
DB	5221	ACTCATCAGAGAGTAACTCGAAGTCTCTCCCAAGCAGCAGCTTCACTAAGGCTCG	5280
QY	1515	ValLeuProValGlyThrGlnLeuProAlaGlyThrLeuProSerGlnGlnLeuProPro	1534
DB	5281	GTGCTGCAAGTGGTACTGAATCTCCAGAGTACTTCAACCAAGCAGCAGCTGCACT	5340
QY	1535	PheProGlyProSerLeuThrGlnSerGlnGlnProLeuGluAspLeuAspAlaGlnLeu	1554
DB	5341	TTTCCAGACCTTCTCAACCACTCCAGCACTTCAAGGATCTTGAAGCTCAATTG	5400
QY	1555	ArgArgThrLeuSerProGlu**IleThrValThrSerAlaValGlyProValSerMet	1574
DB	5401	AGAGAACTTATGTCAGAGATTAATCAAGTACTTGGCGTTGCTGCTGCTGCTGCTG	5460
QY	1575	AlaAlaProThrAlaIleThrGlnAlaGlyThrGlnProGlnGlyValSerGlnVal	1594
DB	5461	GGGCTCCAAAGCAATCAGAGACAGAGACAGAGCCTCAGAGAGGTTGTTCTCAAGTC	5520
QY	1595	LysGlnGlyProValLeuAlaThrSerSerGlyAlaGlyValPheLeuMetGlyArgPhe	1614
DB	5521	AAAGAGGCTGCTCTCAAGCACTTCAAGAGCTGAGTGTATTAAGATGGAGATTT	5580
QY	1615	GlnValSerValAlaAlaAspGlyAlaGlnIleGlnGlyIleAsnIleSerGlnAspAla	1634
DB	5581	CAGTTTCTGTTGACGACAGCGGTGCCAGAAAGGGTAAATAATCAAGAAATGCA	5640
QY	1635	LysSerValHisPheGluSerSerThrSerGluSerSerValLeuSerSerSerPro	1654
DB	5641	AAGCTGTTCATTTGAATCCAGCAGCTCAGAGTCTCAAGTCTATCAAGTAGTGTCA	5700
QY	1655	GluSerThrLeuValValProGluProGlnGlyIleThrIleProGlyIleSerSerAsp	1674
DB	5701	GAGAGTACTTGTGTAACCAAGAGCCGATGCAATACATCTCTGATATCTTCAAGT	5760
QY	1675	ValProGluSerAlaHisIleThrThrAlaSerGlnAlaIleSerSerAspThrGlyGlnPro	1694
DB	5761	GTCAGAGAGTGGCCCAAAATCTACTGCTCAGAGGAGCAAGTCAAGTCACTGGGAGCT	5820
QY	1695	ThrIleValGlyArgPheGlnValThrThrThrAlaAsnIleValGlyArgPheSerVal	1714

Db 5821 ACCAAGGTGGACGTTTCAAGTGAACACTACAGCAACAAAGTGGCTTCTCTGTA 5880
 Qy 1715 SerLySthGluAspLysIleThrAspThrLysLysGluGlyProValAlaSerPro 1734
 Db 5881 TCAAAACAGAGAGACAGATCACTGACACAAAGAAAGAAAGACAGTGCATCTCTCT 5940
 Qy 1735 PheMetAspLeuGluGlnAlaValLeuProAlaValIleProLysLysGluLysProGlu 1754
 Db 5941 TTTATGATTTGGAAACAAGCTTCTCTCTGATGATACCAAGAAAGAAAGAGCTGAA 6000
 Qy 1755 LeuSerGluProSerHisLeuAsnGlyProSerSerAspProGluAlaAlaPheLeuSer 1774
 Db 6001 CTGTCAAGACCTTCACTCACTTAATGAGCGCTTCTGACCCCGAGCGCGCTTTTAAAGT 6060
 Qy 1775 ArgAspValAspAspGlySerGlySerProHisSerProHisLeuSerSerLysSer 1794
 Db 6061 AGGATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 6120
 Qy 1795 LeuProSerGlnAsnLeuSerGlnSerLeuSerAsnSerPheAsnSerSerLysSer 1814
 Db 6121 CTTCCTACCCCAAGATCTAAGTCAAGCTTAACTTAATTAATTAATTAATTAATTAATTA 6180
 Qy 1815 SerAspAsnGluSerAspIleGluAspGluAspLeuLysLeuLysLeuLysLeuLys 1834
 Db 6181 AGCGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 6240
 Qy 1835 AspLysHisLeuLysGluIleGlnAspLeuGlnSerAspGlnLysHisGluIleGlnSer 1854
 Db 6241 GATAAACAATCTCAAAAGAGATTCAGAGACTGACAGAGTCCCGAAGACATGAATTTGAATCT 6300
 Qy 1855 LeuTyThrLysLysLeuLysValProProAlaValIleIleProProAlaAlaProLeu 1874
 Db 6301 TTGTATACCAAACTGGGCAAGTGGCCCTGCTGTTATTAATTCCTCCCGAGCTGCCCTT 6360
 Qy 1875 SerGlyArgArgArgArgProThrLysSerLysGlySerLysSerSerArgSerSerSer 1894
 Db 6361 TCGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6420
 Qy 1895 LeuGlyAsnLysSerProGlnLeuSerGlyAsnLeuSerGlyGlnSerAlaAlaSerVal 1914
 Db 6421 TTGGGGAATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6480
 Qy 1915 LeuHisProGlnGlnThrLeuHisProProGlyAsnIleProGluSerGlyGlnAsnGln 1934
 Db 6481 TTGCACCCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6540
 Qy 1935 LeuLeuGlnProLeuLysProSerProSerSerAspAsnLeuTyThrAlaPheThrSer 1954
 Db 6541 CTGTTTCAAGCCCTTAAGCAATCTCCCTCCAGTACAACTTAATTAATTAATTAATTAATTA 6600
 Qy 1955 AspGlyAlaIleSerValProSerLeuSerAlaProGlyGlnGlyLysThrSerSerThrAsn 1974
 Db 6601 GATGTGTCATTTCAAGTACCAAGCTTTCTCTCAAGTCAAGAGAACAGAGAGAGAGAGAGAG 6660
 Qy 1975 ThrValGlyAlaThrValAsnSerGlnAlaAlaGlnAlaGlnProProAlaMetThrSer 1994
 Db 6661 ACTGTTGGGGCAACAGTGAACAGCCCAAGCCCAAGCTGACCTCTCCGCAATGACGTCC 6720
 Qy 1995 SerArgLysGlyThrPheThrAspAspLeuHisLysLysLeuValAspAsnTPPAlaArgAsp 2014
 Db 6721 AGCAGGAAGGAGCAATTCACAGATGACTTGACAAAGTTGGTGAACAATTTGGGGCCGAGAT 6780
 Qy 2015 AlMetAsnLeuSerGlyArgArgGlySerLysGlyHisMetAsnTyrgLysGlyProGly 2034
 Db 6781 GCCATGAATCTCTCAGGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6840
 Qy 2035 MetAlaArgLysPheSerAlaProGlyGlnLeuCysIleSerMetThrSerAsnLeuGly 2054
 Db 6841 ATGCAAGAGAAATCTCTGCACTGGGCAACTGTGATCTCCAGTCACTGCAAGCTGGGT 6900
 Qy 2055 GlySerLysProLysSerAlaAlaSerAlaThrSerLeuGlyHisPheThrLysSerMet 2074

Db 6901 GGCCTGCCCCCATCTCTGACGATCAGTACTCTCTAGTCACTTCAACAGTCTAG 6960
 Qy 2075 CysProProGlnGlnTyrgLysPheProAlaThrProPheGlyAlaGlnTyrSerGlyThr 2094
 Db 6961 TGCCCCCAAGAGATAGAGCTTTTCACACTACCTCCATTTGGAGCTCAATGAGTGGAGAG 7020
 Qy 2095 GlyGlyProAlaProGlnProLeuGlyGlnPheGlnProValGlyThrAlaSerLeuGln 2114
 Db 7021 GGTGGCCAGACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7080
 Qy 2115 AsnPheAsnIleSerAsnLeuGlnLysSerIleSerAsnProProGlySerAsnLeuArg 2134
 Db 7081 AATTTCACATCAGCAATTTGCAAGAAATCCATCAGAACCCCGAGCTCCAACTGGGG 7140
 Qy 2135 ThrThr 2136
 Db 7141 ACCACT 7146
 RESULT 6
 AR302602 6327 bp DNA linear PAT 12-JUN-2003
 LOCUS AR302602
 DEFINITION Sequence 29 from patent US 6541252.
 ACCESSION AR302602
 VERSION AR302602.1 GI:31690900
 KEYWORDS
 SOURCE
 ORGANISM
 Unknown.
 REFERENCE
 1 (bases 1 to 6327)
 AUTHORS Walke,D.W., Hilpun,E., Donoho,G. and Turner,C.A. Jr.
 TITLE Human kinases and polynucleotides encoding the same
 JOURNAL Patent: US 6541252-A 29 01-APR-2003;
 FEATURES
 source
 1..6327
 /organism="unknown"
 /mol_type="genomic DNA"
 ORIGIN
 Alignment Scores:
 Pred. No.: 5.97e-254 Length: 6327
 Score: 10646.00 Matches: 2108
 Percent Similarity: 98.69% Conservative: 0
 Best Local Similarity: 98.69% Mismatches: 0
 Query Match: 98.46% Indels: 28
 DB: 6 Gaps: 1
 US-10-010-720-14 (1-2136) x AR302602 (1-6327)
 Qy 1 MetSerGlyGlyAlaAlaGluLysGlnSerSerThrProGlySerLeuPheLeuSerPro 20
 Db 1 ATGTCTGGCGGCGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 60
 Qy 21 ProAlaProAlaProLysAsnGlySerSerSerAspSerSerValGlyLysLysGly 40
 Db 61 CCGGCTCTGCCCCCAAGATGAGCTCAAGTCCGATTCCTCGTGGGGAGAAACTGGGA 120
 Qy 41 AlAlaAlaAlaAspAlaValThrGlyArgThrGluGlyLysArgArgArgHisThr 60
 Db 121 GCGGCGGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180
 Qy 61 MetAspLysAspSerArgGlyAlaAlaAlaAlaThrThrThrThrThrGluHisArgPhePhe 80
 Db 181 ATGAGCAAGAGACAGCCCTGGGGGGGGCGGCAACACACACACTGAGCACCGCTTCTTC 240
 Qy 81 ArgArgSerValIleCysAspSerAsnAlaThrAlaLeuGluLeuProGlyLeuProLeu 100
 Db 241 CGCCGAGCGTCACTGCGACTCCAAATGCACATGACATGAGAGCTTCCGCGCTTCCCTT 300
 Qy 101 SerLeuProGlnProSerIleProAlaAlaValProGlnSerAlaProProGluProHis 120
 Db 301 TCCCTGCCCAAGCCAGCATCCCGGCGCTGTCCGAGAGTGTCTCCACCGAGAGCCCAAC 360
 Qy 121 ArgGluGluThrValThrAlaThrAlaThrSerGlnValAlaGlnGlnProProAlaAla 140

Qy	861	ArgSerArgSerArgHisGluLysThrSerArgProLysLeuArgIleLeuAsnValSer	880
Db	2497	AGAGATCGCTCTCGACATGAAAAAACTTCACGCCCAAAATTAAGATTTTGAATGTTTCA	2556
Qy	881	AsnLysGlyAspArgValValGluCysGlnLeuGluThrHisAsnArgLysMetValThr	900
Db	2557	MTTAAAGGACCGAGTAGTAAGTAATGTCAATTAGACTCATTAATAGAAAATGTGTTCA	2616
Qy	901	PhelLysPheAspLeuAspGlyAspAsnProGluGluIleAlaThrIleMetValAsnAsn	920
Db	2617	TTCAATTTGACCTGATGTGTACACCCCGAGAGATAGCAAAATTAATGTGTAAACAT	2676
Qy	921	AspPheIleLeuAlaIleGluArgGlnSerPheValAspGlnValArgGluIleIleGlu	940
Db	2677	GACTTATTTCTAGCAATAGAGAGAGAGTGGTTGTGATCAAGTCGAGAAATTAATGAA	2736
Qy	941	LysAlaAspGluMetLeuSerGlnAspValSerValGluProGluGlyAspGlnGlyLeu	960
Db	2737	AAAGCTGATGAAATGCTCAGTAGAGATGTCAGTGTGAAACCAAGGGTGATCAGAGGATG	2796
Qy	961	GluSerLeuGlnGlyLysAspAspTyrGlyPheSerGlySerGlnLysLeuGluGlyGlu	980
Db	2797	GAGAGTCAACAGAGAAAGATGACTATGGCTTTTCAGGTTCTCAGAAATTTGAAAGAGAG	2856
Qy	981	PhelLysGlnProIleProAlaSerSerMetProGlnGlnIleGlyIleProThrSerSer	1000
Db	2857	TTCAAAACCAACAAATTCCTCGCTCTTCATGCCACAGCAAAATAGGACTTCCTACCACTT	2916
Qy	1001	LeuThrGlnValIleHisSerAlaGlyArgArgPheIleValSerProValProGluSer	1020
Db	2917	TTAACTCAAGTTGTCATTCGCGGGAAGCGGTTTATAGTAGAGTCCGTGCCAGAAAGC	2976
Qy	1021	ArgLeuArgGluSerLysValPheProSerGluIleThrAspThrValAlaAlaSerThr	1040
Db	2977	CGATTACAGAAATCAAAAGTTTCCCGCAGTAATTAACAAATACAGTTGCTGCTCCTACA	3036
Qy	1041	AlaGlnSerProGlyMetAsnLeuSerHisSerAlaSerSerLeuSerLeuGlnAla	1060
Db	3037	GCTAGAGACCTCGAATGAACTTGTCTCCTCATCTGCATCATCCCTTAGCTCAACAAGCC	3096
Qy	1061	PheSerGluLeuArgArgAlaGlnMetThrGluGlyProAsnThrAlaProProAsnPhe	1080
Db	3097	TTTCTGAACTTAAACGTGCCAAATAGCAAGAGACCCAAVACAGCACTCCCAAACTTT	3156
Qy	1081	SerHisThrGlyProThrPheProValValProProPheLeuSerSerIleAlaGlyVal	1100
Db	3157	AGTATACAGAACCAACATTTCCAGTAGTACCTCTTCTTAAGTAGCATTTGCTGGAGTC	3216
Qy	1101	ProThrThrAlaAlaAlaThrAlaProValProAlaThrSerSerProProAsnAspIle	1120
Db	3217	CCAACCAAGACAGACCAAGCACAGCAAGTCCCTGCACAACAAGCAAGCCCTCCAAATGACCT	3276
Qy	1121	SerThrSerValIleGlnSerGlnValThrValProThrGluGluGlyIleAlaGlyVal	1140
Db	3277	TCCACATCAAGTAATTCAGTCTGAGGTTACATGTGCCACTGAAGGGGATTTGCTGGAGTT	3336
Qy	1141	AlaThrSerThrGlyValValThrSerGlyLysLeuProIleProProValSerGluSer	1160
Db	3337	GCCACCAAGCACAGGTGTGTACTTCAAGTGTCTCCCATACCACTGTGTCTGAATCA	3396
Qy	1161	ProValLeuSerSerValValSerSerIleThrIleProAlaValValSerIleSerThr	1180
Db	3397	CCAATACCTTCCAGCGTAGTTTCAAGTATCACATATACCTGCAAGTTGTCTCAATATCTACT	3456
Qy	1181	ThrSerProSerLeuGlnValProThrSerThrSerGluIleValValSerSerThrAla	1200
Db	3457	ACATCCCGCTACTTCAAGTCCCAACATCCACATCTAGAGATCGTTGTTTCTAGTAGCA	3516
Qy	1201	LeuThrProSerValThrValSerAlaThrSerAlaSerAlaGlyLysThrAlaThr	1220
Db	3517	CTGATCTCTTCAAGTAAGTTTACGCAACTTCAGCCCTCTGAGGGGGCAGTACTGCTACC	3576

Qy	1221	ProGlyProLysProProAlaValValSerGlnGlnAlaAlaGlySerThrThrValGly	1240
Db	3577	CCAGTCTTAAGCCTCCAGCTGTAGTATCTACAGACGACGACGACACTACTGTGGGA	3636
Qy	1241	AlaThrLeuThrSerSerValSerThrThrThrSerPheProSerThrAlaSerGlnLeuSer	1260
Db	3637	GCCACATTAAATCATGATTTCTACACCACTTCACTTCCCAAGACAGCTTCCAGAGCTGTC	3696
Qy	1261	IleGlnLeuSerSerSerSerThrSerThrProThrLeuAlaGluThrValValValSerAla	1280
Db	3697	ATTCAAGCTTACAGACAGTACTTCTTACTTCACTTGAAGCTGTAACCGTGTAGTTCGCCA	3756
Qy	1281	HisSerLeuAspLysThrSerHisSerSerThrThrGlyLeuAlaPheSerLeuSerAla	1300
Db	3757	CACCTACAGTAATAGACATCTCATAGCACTCACTGATGGCTTTCTCCCTCTCTCCA	3816
Qy	1301	ProSerSerSerSerSerSerProGlyValAlaGlyValSerSerTyrIleSerGlnProGly	1320
Db	3817	CCATCTTCTCTCTCTCTCTCGAGACGAGAGTGTACTATTAATTTCTCAGCTGTGGG	3876
Qy	1321	LeuHisProLeuValIleProSerValIleAlaSerThrProIleLeuProGlnAlaAla	1340
Db	3877	CTGCATCTTGTGTCATTCATTCAGTAAATGCTTCTCTCTATTTCTTCCCAAGACACA	3936
Qy	1341	GlyProThrSerThrProLeuLeuProGlnValProSerIleProProLeuValGlnPro	1360
Db	3937	GGACCTACTTCTACACCTTATTAATCCCAAGTACTAGTATCCACCTTGTGTACAGCT	3996
Qy	1361	ValAlaAsnValProAlaValGlnGlnThrLeuIleHisSerGlnProGlnProAlaLeu	1380
Db	3997	GTTCACATGCGCTCGTAGACAGCACTAATTCATAGTCAAGCTCAACACAGCTTTG	4056
Qy	1381	LeuProAsnGlnProHisSerHisCysProGluValAspSerAspThrGlnProLysAla	1400
Db	4057	CTTCCCAACCAAGCCCACTACTGCTCTGTAATTAATTCATTAATTCATTAATTCATTA	4116
Qy	1401	ProGlyIleAspAspIleLysThrLeuGluGluLysLeuAspSerLeuPheSerGluHis	1420
Db	4117	CTTGGAATGATGACATTAAGACATCTAGAACAAAGCTGCGTCTCTGTTCAAGTAAC	4176
Qy	1421	SerSerSerGlyAlaGlnHisAlaSerValSerLeuGluThrSerLeuValIleGluSer	1440
Db	4177	AGCTCATGTGAGTCTCAGATGCTCTGTCTCACTGAGACCTCACTAGTATAGAGAGC	4236
Qy	1441	ThrValThrProGlyIleProThrThrAlaValAlaProSerLysLeuLeuThrSerThr	1460
Db	4237	ACTGTCAACCAAGGACATCCAACTACTGCTGTGCACCAAGCAAACTCTGACTTCTAAC	4296
Qy	1461	ThrSerThrCysLeuProProThrAsnLeuProLeuGlyThrValAlaLeuProValThr	1480
Db	4297	ACAAGTACTTGTCTTACCAACCAATTAACCACTTAAGAACTGTGCTTCCAGTTACA	4356
Qy	1481	ProValValThrProGlyGlnValSerThrProValSerThrThrThrSerGlyValLys	1500
Db	4357	CCAGTGTCTACACTGTGGCAAGTTTCAACCCAGTCAAGCACTACTACATCAGAGAGTAAA	4416
Qy	1501	ProGlyThrAlaProSerLysProProLeuThrLysAlaProValLeuProValGlyThr	1520
Db	4417	CTTGAACTGTCTCTCCCAAGCCACTTAACCTTAAGGCTCGGGCTGCCAGAGGGGACT	4476
Qy	1521	GluLeuProAlaGlyThrLeuProSerGluGlnLeuProProPheProGlyProSerLeu	1540
Db	4477	GAACCTTCAGACAGTACTTACCAAGGAGACAGTGCACCTTTTCCAGGACCTTCTCTA	4536
Qy	1541	ThrGlnSerGlnGlnProLeuGluAspLeuAspAlaGlnLeuArgArgThrLeuSerPro	1560
Db	4537	ACCCAGTCCCAAGCAACTTCAAGAGATCTGAAGCTCAATTTGAAGAACACATTAGTCCA	4596
Qy	1561	Glu***IleThrValThrSerAlaValGlyProValSerMetAlaAlaProThrAlaIle	1580
Db	4597	GAGATATACAGTACTTGTGCGGTGTGCTGTGTGTCCATGCGGGCTCCAAACAGCAATC	4656
Qy	1581	ThrGluAlaGlyThrGlnProGlnLysGlyValSerGlnValLysGluGlyProValLeu	1600

```

Db 4557 ACAGAGCAGGAAACACACCTCAGAGGGTCTTCTCAGCTCAAGCAAGAGGCCCTGTCTTA 4716
QY 1601 AAlaThrSerSerGlyAlaGlyValPheIysMetGlyArgPheGlnValSerValAlaAla 1620
Db 4717 GCACTAGTTCAGAGAGCTGTGTCTTTTAAAGTGGGACGATTCAGGTTCTGTGTGACGCA 4776
QY 1621 AspGlyAlaGlnIlyGlnGlyIlyLeuAsnIlySerGlnAlaIlySerValHisPheGln 1640
Db 4777 GACGGTGCCCAAGAAAGAGGTAATAATAGTCAGAAAGATGCAAAAGTCGTTCATTTTGA 4836
QY 1641 SerSerThrSerGlnSerSerValIleuSerSerSerSerProGlnSerThrIleuValIys 1660
Db 4837 TCCAGCACTCAGAGCTCCTCAGTGTCTATCAAGTAGTACTCCAGAGAGTACTTGTGAA 4896
QY 1661 ProGlnProAsnGlyIleThrIleProGlyIleSerSerAspValProGlnSerAlaHis 1680
Db 4897 CCAGAGCCGAATGGATACCATCCCTGGTATCTCTTCAGATGTGCCAGAGAGTGCAC 4956
QY 1681 LysThrThrAlaSerGlnAlaIlySerAspThrGlyGlnProThrIlyValGlyArgPhe 1700
Db 4957 AAACCTACTGCTCAGAGGCAAGTCAAGCACTGGGAGCCTTCCAGAGTTGGACGTTT 5016
QY 1701 GlnValThrThrAlaAsnIlyValGlyArgPheSerValSerIysThrGlnAspIys 1720
Db 5017 CAGGTGACACTACGCAACCAAGTGGTCTGTTCTGTATCAAAACAGAGACAG 5076
QY 1721 IleThrAspThrIlyIlyGlnGlyProValAlaSerProPheMetAspLeuGln 1740
Db 5077 ATCCTGACAGCAAAAGAAAGAGACCAAGTGCATCTCTCTTATGATTTGGACAA 5136
QY 1741 AlaValIleuProAlaValIleProIlyIysGlnIlySerProGlnIleuSerGlnProSerHis 1760
Db 5137 GCTGTTCTTCTCTGCTGATACCAAGAAAGAGCCTGAACCTGCAAGCTTCACAT 5196
QY 1761 LeuAsnGlyProSerSerAspProGlnAlaIlePheIleuSerArgAspValAspAspGly 1780
Db 5197 CTAAATGGGCCGTCTTTCGACCCGAGGCCGCTTTTAAAGAGGAGATGTGATATGAT 5256
QY 1781 SerGlySerProIleSerProIleGlnIleuSerIlySerIleuProSerGlnAsnIleu 1800
Db 5257 TCCGATACCTCACACTCCGCCCATAGCTGAGCTCAAGAGGCTTCCACAGATCA 5316
QY 1801 SerGlnSerIleuSerAsnSerPheAsnSerSerIlyMetSerSerAspAsnGlnSerAsp 1820
Db 5317 AGTCAAAGCCTTATGATTAATTCATTTACTCTCTTACATGAGTACGCAAAAGATCAAT 5376
QY 1821 IleGlnAspGlnAspLeuIlyLeuGlnIleuArgIleuArgLeuArgAspIlyHisIleuIysGln 1840
Db 5377 ATCGAAGATGAAGACTTAAAGTTAGAGTGGCAGCAGACTACGATTAACATCTCAAGAG 5436
QY 1841 IleGlnAspLeuGlnSerArgGlnIlyHisIleuIleuSerIleuIlyThrIlyIleuGly 1860
Db 5437 ATTCAGGACCTGGCAGAGTCCGCAAGAGATGAATTTGATCTTTTATTAACCAAACTGGGC 5496
QY 1861 LysValIleuProAlaValIleIleProProAlaAlaIleProLeuSerGlyArgArgArgArg 1880
Db 5497 AAGGTGCCCTCGCTGTATTTATTTCCCCAGCTGCTCCCTTTCCAGGAGAAAGAGACGA 5556
QY 1881 ProThrIlySerIlyGlySerIlySerSerArgSerSerSerSerIleuGlyAsnIlySerPro 1900
Db 5557 CCCACTAAAGAGCAAGGACGCAAAATCTAGTCGAACAGATCTCTTGCGGAAATTAAGCCCC 5616
QY 1901 GlnIleuSerGlyAsnIleuSerGlyGlnSerAlaAlaSerValIleuHisIleuProGlnIleuThr 1920
Db 5617 CAGCTTTCAAGTAACTGCTGTGGTCAAGTGCAGCTTCAAGTTCACCCCCAGACGACC 5676
QY 1921 LeuHisProProGlyAlaAsnIleProGlnSerGlyGlnAsnGlnIleuLeuGlnIleuProIlyHis 1940
Db 5677 CTCACACCTCTCTGGCAACATCCCAAGTCCGAGGCAAGATCAGCTTTACAGCCCTTTAAG 5736
QY 1941 ProSerProSerSerAspAsnIleuIlySerAlaPheThrSerAspGlyAlaAlaIleSerVal 1960

```

```

Db 5737 CCATCTCCCTCCAGTACCAACCTCTATTCAGCTTACCAAGTGAATGTGTCATTTCACTA 5796
QY 1961 ProSerIleuSerAlaProGlyGlnGlyIlyThrSerSerThrAsnThrValGlyAlaThrVal 1980
Db 5797 CCAAGCCTTTCTGTCTCAGAGTCAAGGAAACGAGCAACCAACACTGTTGGGCAACAGTG 5856
QY 1981 AsnSerGlnAlaIleGlnIleGlnProProAlaMetThrSerSerArgIlyThrPhe 2000
Db 5957 AACAGCAACGCCGCCCAAGCTCAGCTCCTCCCTACAGTCCAGCAGAAAGGACATTC 5916
QY 2001 ThrAspAspLeuHisIlyIleuValAspAsnTrpAlaArgAspAlaMetAsnIleuSerGly 2020
Db 5917 ACAGATGATCTTGCAAGTGTGTAGACATTTGGGCGCCAGAGATGCCATGATCTCTCAGGC 5976
QY 2021 ArgArgGlySerIlyGlyHisIleMetAsnTrpGlnIlyProGlyMetAlaArgIlyPheSer 2040
Db 5977 AGGAGAGAAAGCAAGGACATATTAATTAAGGACCTTGAAATGGCAAGAGATTTCT 6036
QY 2041 AlaProGlyGlnIleuCysIleSerMetThrSerAsnIleuGlyIlySerAlaProIleSer 2060
Db 6037 GCACCTGGGCAACTGTGCATCTCCATGACCTCGAACCTGGGTGGCTGCCCCCATCTCT 6096
QY 2061 AlaAlaSerAlaThrSerIleuGlyHisPheThrIlySerMetCysProProGlnGlnIly 2080
Db 6097 GCACCATCAGCTACTCTCTTCAAGTCACTTCAACCAAGTCTAATGTGCCCCCAGACAGAT 6156
QY 2081 GlyPheProAlaThrProPheGlyAlaGlnIlyPheSerGlyThrGlyIlyProAlaProGln 2100
Db 6157 GGGCTTTCAGCTACCCCAATTTGGCTCATATGAGTGGAGCGGTGGCCCAACACACAG 6216
QY 2101 ProIleuGlyGlnPheGlnProValGlyThrAlaSerLeuGlnAsnPheAsnIleSerAsn 2120
Db 6217 CCACCTGGCCAGTTCCAACCTGTGGAACTGCTCTCTTGAGATTCATCAATCAGCAAT 6276
QY 2121 LeuGlnIlySerIleSerAsnProProGlyIlySerAsnLeuAlaGlnThr 2136
Db 6277 TTGCAGAATCCATCAGCAACCCCGGCTCCCAACCTGCGGACCACT 6324

RESULT 7
AR302599 AR302599 6606 bp DNA linear PAT 12-JUN-2003
LOCUS DEFINITION Sequence 23 from patent US 6541252.
ACCESSION AR302599
VERSION AR302599.1 GI:31690897
KEYWORDS
SOURCE unknown.
ORGANISM unknown.
REFERENCE 1 (bases 1 to 6606)
AUTHORS Walke, D.W., Hilbun, E., Donoho, G. and Turner, C.A. Jr.
TITLE Human kinases and polynucleotides encoding the same
JOURNAL Patent: US 6541252-A 23 01-Apr-2003;
FEATURES
source Location/Qualifiers
1..6606
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Alignment Scores:
Pred. No.: 1,54e-252 Length: 6606
Score: 10589.50 Matches: 2108
Percent Similarity: 94.57% Conservative: 0
Best Local Similarity: 94.57% Mismatches: 0
Query Match: 97.94% Indels: 121
DB: Gaps: 2

US-10-010-720-14 (1-2136) x AR302599 (1-6606)
QY 1 MetSerGlyGlyAlaAlaGlnIlyGlnSerSerThrProGlySerIleuPheLeuSerPro 20
Db 1 ATGTCTGGCGGGGCGGCGAGAGAGACAGACACTCCCGGTTCTCTCTCTCGCGCG 60
QY 21 ProAlaProAlaProIlyAsnGlySerSerSerAspSerSerValGlyIlyIlyLeuGly 40

```

```

Db 61 CCGGCTCTGCCCCCAAGAAATGGCTTCAGCTCCGATCTCCGCTGGGAGAAACCTGGCA 120
Qy 41 AAlaAlaAlaAspAlaValThrGlyArgThrGluGlyArgArgArgArgIleThr 60
Db 121 GCGCGGCGCGCGGAGCTGTGACCGCGAGACCGAGAGATACAGCGCGCGCGCACT 180
Qy 61 MetAspIleAspSerArgIleAlaAlaIleThrThrThrThrThrIleIleArgPhe 80
Db 181 ATGACAAAGACAGCGCTGGGCGCGCGCGCACTTACCACTGAGCACTGAGCACTGCTTTC 240
Qy 81 ArgArgSerValIleCysAspSerAsnAlaThrAlaLeuGluLeuProGlyLeuProLeu 100
Db 241 CGCGGAGCGTCATCTGCGACTCCAGTCCACTGCACTGAGAGCTTCCCGGCTTCTCTT 300
Qy 101 SerLeuProGluProSerIleProAlaAlaValProGluSerAlaProProGluProHis 120
Db 301 TCCGCGCCCAAGCCAGATCCCGCGGCTGTCCGAGAGTCTCCACCGAGCCCGAC 360
Qy 121 ArgGluGluThrValThrAlaThrAlaThrSerGluValAlaGluGluProProAlaAla 140
Db 361 CGGAGAGAGACCGGACCGCACCGCACTTCCAGTAGGCCAGCACTTCCAGCGCT 420
Qy 141 AlaAlaProGluGluGluAlaValAlaGlyProAlaProSerThrValProSerSerThr 160
Db 421 GCGCGCTTGGGAGAACAGCGCGTGGCGGCGCTGCGCTGCGACTGCTCCAGAGTACC 480
Qy 161 SerIleAspArgProValSerGluProSerLeuValGlySerIleGlyGluGluProPro 180
Db 481 AGCAAAAGACCGCCAGGTGTCCAGCTTACCTTGTGGGAGACAAAGAGAGACCGCGCG 540
Qy 181 AlaArgSerGlySerGlyGlyGlySerAlaGlyGluProGluGluGluArgSerGluGlu 200
Db 541 GCGAGAAAGTGCAGCGCGCGCGCACCGCAAGAGAGCACAGAGAGAACGAGCAGAG 600
Qy 201 GluAspAspIleGluGluLeuGluThrIleValAlaGlyMetSerAsnAspGlyArgPhe 220
Db 601 CAGAGATGATCGAAGAGCTGAGAACCAAGCGCTGGGAATGCTTAAAGATGCGCGCTT 660
Qy 221 LeuIleAspPheAspIleGluIleGlyArgGlySerPheIleThrValIleValGlyLeuAsp 240
Db 661 CTCAAGTTTGACATCGAAATCGCAGAGCGCTCTTAAAGACGCTCAAAAGGCTGGAC 720
Qy 241 ThrGluThrThrValGluValAlaIleProCysGluLeuGluAspArgIleThrThrIleSer 260
Db 721 ACTGAAACCAACCGGAAAGTGGCTGGTGAATCGACAGATCGAAAATTAAACAAGTCT 780
Qy 261 GluArgGluArgPheGlyGluGluAlaGluMetLeuIleGlyLeuGluIleIleProAsnIle 280
Db 781 GAGAGGCGAGATTTAAAGAAAGAGTGAATGTAAAGGTCTTCAAGCATCCCAATATT 840
Qy 281 ValArgPheIleAspSerIleProGluSerThrValIleGlyGlyIleValIleValIle 300
Db 841 GTTAGATTTTATGATCTCTGGGAATCCACAGTAAAGAAAGAGTCAATGTTTGGTG 900
Qy 301 ThrGluLeuMetThrSerGlyThrLeuIleThrIleValIleGlyArgPheIleValIleMet 320
Db 901 ACTGAACTTATGAGCTGGAACACTTAAACGATCTGAAAAGGTTTAAAGTAGTAG 960
Qy 321 IleIleValLeuArgSerIleProCysArgGluIleLeuIleGlyLeuGluIleLeuIleThr 340
Db 961 ATCAAAATCTCAAGACTGGTGCCTGCAATCTTAAAGTCTTCAAGTCTTCAATCT 1020
Qy 341 ArgThrProIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIle 360
Db 1021 CGAAGTCACTATCACTACCGGAGCTTAAAGTCAACAATCTTAAACCGGCGCT 1080
Qy 361 ThrGlySerValIleIleGlyAspLeuGlyLeuAlaThrLeuIleArgAlaIlePheAla 380
Db 1081 ACTGGGCTCAGTCAAGATTTGAGAGCTGCTGGCAACCTGGAAGCGGCTTCTTTGCC 1140
Qy 381 LysSerValIleGlyThrProGluPheMetAlaProGluMetIleGluGluIleIleIle 400
Db 1141 AAGATGTGATAGGTACCCAGAGTTCATGCGCCCTGATGATGTAGAGAAATATGAT 1200
Qy 401 GluSerValAspValIleValAlaPheGlyMetCysMetLeuGluMetAlaIleThrSerGluIle 420
Db 1201 GAATTCGTTGACGTTTATGCTTTTGGATGGATGTCATGCTTGAAGTGTCAATCTGAATAT 1260
Qy 421 ProIleSerGluCysGluAsnAlaAlaGluIleIleIleIleIleIleIleIleIleIleIle 440
Db 1261 CTTTACTGGAGTCCCAAAATGCTGCACAGATCTACCGTCCGAGACAGTGGGATGAG 1320
Qy 441 ProIleSerPheAspIleValAlaIleProGluValIleGluIleIleGluIleCysIle 460
Db 1321 CCAGCCAGTTTGAACAAGTGAATTCCTGAGTGAAGAAATATTATGAAGATGATCA 1380
Qy 461 ArgGluAsnIleAspGluArgIleIleIleIleIleIleIleIleIleIleIleIleIle 480
Db 1381 CGACAAACAAAGATGAAAGATATTCATCAACAAACCTTTGAAACCAATCTTCCAA 1440
Qy 481 GluGluThrGlyValArgValGluLeuAlaGluGluAspAspGlyGluIleAlaIle 500
Db 1441 GAGGAAACAGAGATACCGGTAGATTAAGCAAGAAAGATGTGAGAAAATATGACATA 1500
Qy 501 LysLeuIlePheLeuArgIleGluAspIleIleIleIleIleIleIleIleIleIleIle 520
Db 1501 AAATTATGCTACCTATTTGAAGATATTAAAGAAATTTAAAGGAAATATCAAAAGATTAGAA 1560
Qy 521 AlaIleGluPheSerPheAspLeuGluArgAspValProGluAspValAlaGluIleMet 540
Db 1561 GCTATTAGTTTCTTTTGAATTTAGAGAGATGTCCAGAAAGTGTGCAAGAAATG 1620
Qy 541 ValGluSerGlyThrValCysGluGlyAspIleIleIleIleIleIleIleIleIleIleIle 560
Db 1621 GTAGAGTCTGGATATGCTGTGAAGATGATCAAGAACATGCTTAAAGTCAATCAAGAC 1660
Qy 561 ArgValSerIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIle 580
Db 1681 AGATATCAATTAATTAAGAGAAACGAGACAGCGCAGTTGTTGAGGAGAGCAAGAA 1740
Qy 581 LysIleIleGluGluGluGluSerIleIleIleIleIleIleIleIleIleIleIleIle 600
Db 1741 AAAAAAAGAGAGAGAGACAGTCTTAAACAGAGGTAAACAAATCAATGCTTCCAG 1800
Qy 601 ThrGlyIleIleGluLeuProSerIleSerThrGlyIleProThrAlaSerThrThrSer 620
Db 1801 ACGAGATCAAGAGACCTCCCTTCTGCTAGACACCGGCACTACTACTCTTCACTTCA 1860
Qy 621 AlaSerValSerThrGluValGluProGluGluProGluIleAspGluIleGluGlu 640
Db 1861 GCTTCAGTTTCTACAAATGAGAACTGAGAAACCTGAGGACATCAACATCAACAATA 1920
Qy 641 GluThrGluGluProSerIleSerValLeuSerAspGlyThrValAspSerGlyGluGly 660
Db 1921 CAGTACCAAGAACCGATATCTGTATCTGATGAGCGGTGACAGTGTGACAGGGA 1980
Qy 661 SerSerValPheThrGluSerArgValSerSerGluGluIleThrValSerIleGlu 680
Db 1981 TCCCTGCTTCAAGATCTCAGATGAGAGCAGCAACAGCAATTTCAATGTTCCCA 2040
Qy 681 HisGluGluAlaHisSerThrGlyThrValProGlyHisIleProSerThrValGluAla 700
Db 2041 CATGAACAGGACATTTCAAGGACAGATCCAGGCGATATACCTTACTACTGCAAGCA 2100
Qy 701 GluSerGluProHisGlyValIleProProSerSerValAlaGluGluGluSerGluGly 720
Db 2101 CAGTTCAGCCCAATGAGGATATTCACCTCAAGTGTG----- 2139
Qy 721 GluProSerSerSerSerIleThrGlyValSerSerSerGluProIleGluHisProGlu 740
Db 2139 ----- 2139
Qy 741 GluGluGluGluIleGluGluIleThrAlaProProGluGluIleThrValGluIleSerLeuSer 760
Db 2140 ---CAGAGGAGATACAGCAGACAGCCCTCTCAACAGACAGTGTCAATTCATTTCA 2196

```

QY 761 GlnThrSerThrSerSerGluAlaThrThrAlaGlnProValSerGlnProGlnAlaPro 780
Db 2197 CAGACATCAACCTCAGTAGAGGCCATCTGCACAGCCAGAGTCAAGCTCCAGCTCCA 2256
QY 781 GlnAlaLeuProGlnValSerAlaGlyLysGln----- 791
Db 2257 CAAGTCTTGCCTCAAGTATACGCTGGAAAACAGGGCTTCCACCTCCAGCTGCACACAG 2316
QY 791 ----- 791
Db 2317 TACCAGAGATCAATATATGTCCTCTTCCAACTGCTCTGTGTGATCCATCTCT 2376
QY 791 ----- 791
Db 2377 ACAGTCTTACCCCTCCATGCCACAGAGTACGCTACACCTGGTACTTCCACA 2436
QY 791 ----- 791
Db 2437 GTGGTGCAGCCTTATGTGGATCAATCTTTTATGTTCTATGGGTGGTGAAGAGACAG 2496
QY 791 ----- 791
Db 2497 GTTCAAGTGTCCAGCCAGAGGAGTTAGCACAGCCCCACTACATCTCTCCACGAA 2556
QY 792 -----SerThrGlnGlyValSerGlnValAlaProAlaGlnProValAlaVal 807
Db 2557 GCAGTGTGGAGAGTACTCAGAGGAGTCTCTCAGTGTCTCTGCAGAGCCAGTGTGACATA 2616
QY 808 AlaGlnProGlnAlaThrGlnProThrThrLeuAlaSerSerValAspSerAlaHisSer 827
Db 2617 GCAGAGCCCAAGTACACCAAGCCAGCCACTTGGCTTCTCTGTGAGACAGTGCACATCCA 2676
QY 828 AspValAlaSerGlyMetSerAspGlyAsnGlyAsnValProSerSerSerGlyValArgHis 847
Db 2677 GATGTGCTTCAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2736
QY 848 GlnGlyAlaGlnThrThrAlaGlnHisGlyThrAlaGlnSerValArgSerSerSerGlyGlu 867
Db 2737 GAGGAGAAACATCAAAACGAGATTAACGAAATCTGTGAGAGTGGCTCTGCACATGAA 2796
QY 868 LysThrSerArgProLysLeuAlaArgLysLeuAlaSerAsnLysGlyAspArgValVal 887
Db 2797 AAACTTACCCGCCAAATTAAGATTTTGAATGTTTCAATTAAGAGACCGAGATATA 2856
QY 888 GlnCysGlnLeuGlyThrHisAsnArgLysMetValThrPheLysPheAspLeuAspGly 907
Db 2857 GAATGTCAATTAGAGACTCATTAATAGAAAATGCTTACATTCAAATTTGACCTACATGCT 2916
QY 908 AspAsnProGlnGluLysAlaThrLysMetValAsnAsnAspPheLysLeuAlaLysGlu 927
Db 2917 GACAAACCCGAGAGAAATAGCAACATTTATGTAACATGACTTATTTCTTACCAATTAAG 2976
QY 928 ArgGlnSerPheValAspGlnValArgGluLysLysLysAlaAspGluMetLeuSer 947
Db 2977 AGAGAGTGGTTGTGATCAAGTGGAGAAATTAATTAAGAAAAGCGAGAAATGCTCAGT 3036
QY 948 GlnAspValSerValGlnProGlnGlyAspGlnGlyLeuGlnSerLeuGlnGlyLysAsp 967
Db 3037 GAGGATGTCAAGTGTGAAACCAAGGCTGATTCAGGATTCGAGAGCTTCACAAAGAAAGAT 3096
QY 968 AspTyrGlnPheSerSerGlySerGlnLysLeuGlnGlyLysPheLysGlnProLysProLys 987
Db 3097 GACTATGGCTTTTACGTTCTCAGAAATTTGAAAGAGAGTTCACAAACCAATTTCTCTGG 3156
QY 988 SerSerMetProGlnGlnLysGlyLysProThrSerSerLeuThrGlnValValHisSer 1007
Db 3157 TCTTCCATGCCACAGAAATAGGACTTCTTACAGTCTTTTAACTCAAGTGTCTTATCT 3216
QY 1008 AlaGlyAlaGlnPheLysValSerProValProGlnSerThrLeuAlaGlnSerLysVal 1027
Db 3217 GCGGAGAGGCGGTTTATGATGAGTCTGTGCCAGAAAGCCGATTCAGAAATCAAAAGTT 3276

QY 1028 PheProSerGlnLysLeuThrAspThrValAlaAlaSerThrAlaGlnSerProGlyMetAsn 1047
Db 3277 TTCCCGCTGAATATACAGATACAGTGTGCTGCTCTTACAGCTCAGAGCCCTGGAAATGAC 3336
QY 1048 LeuSerHisSerAlaSerSerLeuSerLeuGlnGlnAlaPheSerGlnLeuArgArgAla 1067
Db 3337 TTGCTCAGCTGTGATCATCCCTTAGCTTACAAAGAGCCCTTTCTGAACTTAGACGTGCC 3396
QY 1068 GlnMetThrGlnGlyProAsnThrAlaProProAsnPheSerHisGlnGlyProThrPhe 1087
Db 3397 CAATATGACAGAAAGACCAAAACAGACCTTCAAACTTATGATCATACAGAGCAACATTT 3456
QY 1088 ProValValProProPheLeuSerSerLeuAlaGlyValProThrThrAlaAlaAlaThr 1107
Db 3457 CCAGTATGACTCTCTTTCTTATAGATATGCTGAGTCCAGACACAGAGAGCCACA 3516
QY 1108 AlaProValProAlaThrSerSerProProAsnAspLysSerThrSerValLysGlnSer 1127
Db 3517 GCACAGTCCCTGCACAAAGACAGCCCTCTAATGACATTTCCACATCAGTAAATCAGTCT 3576
QY 1128 GlnValThrValProThrGlnGlnGlyLysAlaGlyValAlaThrSerThrGlyValVal 1147
Db 3577 GAGGTACAGTGCCTCAGTGAAGAGGAGATGCTGAGTGCACACAGCAGAGTGTGTA 3636
QY 1148 ThrSerGlyGlyLeuProLysProProValSerGlnSerProValLeuSerSerValVal 1167
Db 3637 ACTTCAAGTGTCTCCCATACCACTGTGTCTGATCAGATCAGATCCTTCCAGCTACTT 3696
QY 1168 SerSerLysThrLysProAlaValAlaSerLysSerThrThrSerProSerLeuGlnVal 1187
Db 3697 TCAAGTATCAAAATACAGCAGTGTCTCAATATCTAATATCCCTCACTTCAACTGC 3756
QY 1188 ProThrSerThrSerGlnLysValValSerSerThrAlaLeuGlyProSerValThrVal 1207
Db 3757 CCGCATTCACATCTAGATGATGTTGTTCTAGTACAGACGTATCTCTTCACTAGTAACT 3816
QY 1208 SerAlaThrSerAlaSerAlaGlyLysSerThrAlaThrProGlyProLysProAla 1227
Db 3817 TCAGCACTTACAGCTCTGCAGAGGAGAGTCTGATCCAGAGTCTTACCTCCAGCT 3876
QY 1228 ValValSerGlnGlnAlaAlaGlySerThrThrValGlyAlaThrLeuThrSerValSer 1247
Db 3877 GTAGATCTCAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGTCT 3936
QY 1248 ThrThrThrSerPheProSerThrAlaSerGlnLeuSerLysLysLysSerSerThr 1267
Db 3937 ACCACCACTTATCTCCAAAGCAGCTTCAACAGCTGATTCATTCAGTACAGCAGTCT 3996
QY 1268 SerThrProThrLeuAlaGluThrValValAlaSerAlaHisSerLeuAspLysThrSer 1287
Db 3997 TCTACTCTTACTTATAGTGAACCGGTGTAGTGGCAGCACTCACTGATTAAGACATCT 4056
QY 1288 HisSerSerThrThrGlyLeuAlaPheSerLeuSerAlaProSerSerSerSerPro 1307
Db 4057 CATAGCAGTACAACTGATGATGCTTCTCCCTCTGACCAATCTTCTCTCTCTCTCT 4116
QY 1308 GlnAlaGlyValSerSerTyrLysGlnProGlyGlyLeuHisProLeuValLysPro 1327
Db 4117 GAGCAGAGAGTGTCTAGTATATTTCTCAGCCTGTGTGGTGCATCTTGGTCTATCCA 4176
QY 1328 SerValLysAlaSerThrProLysLeuProGlnAlaAlaGlyProThrSerThrProLeu 1347
Db 4177 TCAGTATAGTCTTACTACTTCTTCTTCCCAAGCAGCAGGACCTTACTTACACCTTTA 4236
QY 1348 LeuProGlnValProSerLysProProLeuValGlnProValAlaAsnValProAlaVal 1367
Db 4237 TTACCCCAAGTACCTAGTATCCACCCCTGTGACGCTGTGTCACAAATGCTGCTGTA 4296
QY 1368 GlnGlnThrLeuLysHisSerGlnProGlnProAlaLeuLeuProAsnGlnProHisThr 1387
Db 4297 CAGCAGACATTAATCATATGATGAGCTTCAACAGCTTGTCTTCCAAACAGCCCAATCT 4356
QY 1388 HisCysProGlnValAspSerAspThrGlnProLysAlaProGlyLysAspAspLysLys 1407

```
Db 4357 CATTTGCTCGAAGTGAATTTCTGATACAAACCAAGCTCTCGAATGATACATAAG 4416
Qy 1408 ThrLeuGluGluLysLeuArgSerLeuPheSerGluHisSerSerSerGlyValIleHis 1427
Db 4417 ACTTAGAAGAAAGAGCTGGGCTCTGTTCACTGAAACACAGTCACTTGGAGCTAGAGT 4476
Qy 1448 AlaSerValSerLeuGluThiSerLeuValIleGluSerThiValThiProGlyLeuPro 1447
Db 4477 GCGCTGCTCTCAGAGAGCTCAGTACATAGAGAGACAGTCAACACAGCATCCCA 4536
Qy 1448 ThrThrAlaValAlaProSerLysLeuThiSerThiThiSerThiCysLeuProPro 1467
Db 4537 ACTACTGCTGTTGCAACAGCAAACTCTGACTTGTACCAAGACTGCTTACACCA 4596
Qy 1468 ThrAsnLeuProLeuGlyThiValAlaLeuProValThiProValIleThiProGlyGln 1487
Db 4597 ACCAATTTACCACTAGGAAGAGTCTGTTGCAAGTTACCAAGTGTACACCTGGGCA 4656
Qy 1488 ValSerThiProValSerThiThiThiSerGlyValLysProGlyThiAlaProSerLys 1507
Db 4657 GTTCTACCCCACTGACGACTACTACATCAGGAGGAACTGGAACTGGAACTGCTCCCAAG 4716
Qy 1508 ProProLeuThiLysAlaProValLeuProValGlyThiThiGluLeuProAlaGlyThiLeu 1527
Db 4717 CCACTTACCTAAAGGCTCCGGTGTGCTGCAAGTGGTACTGAACTTCCAGCAGTACTCTA 4776
Qy 1528 ProSerGluGlnLeuProProPheProGlyProSerLeuThiGlnSerGlnProLeu 1547
Db 4777 CCCAGCGAGCAGCTGCCACTTTTCCAGAGCTTCTCTAACCCAGTCCCAAGCACTCTTA 4836
Qy 1548 GluAspLeuAspAlaGlnLeuArgArgThiLeuSerProGlu***IleThiValThiSer 1567
Db 4837 GAGATCTTGTGCTCAATTAGAGAAACACTTAGTCCAGAGATATATACAGTACTCTT 4896
Qy 1568 AlaValGlyProValSerMetAlaAlaProThiAlaIleThiGluAlaGlyThiGlnPro 1587
Db 4897 GCGGTGGTCTGCTGCTCAGTGGCTCCACAGCAATCACAGCAAGCAAGCAACAGCTT 4956
Qy 1588 GlnLysGlyValSerGlnValLysGluGlyProValLeuAlaThiSerSerGlyValAlaGly 1607
Db 4957 CAGAAAGGCTGTTTCTCAAGTCAAAAGAGGCCCTGCTCCAGCACTAGTCCAGGAGCTGT 5016
Qy 1608 ValPheLysMetGlyArgPheGlnValSerValAlaAlaAspGlyAlaGlnLysGluGly 1627
Db 5017 GTTTTAAAGATGGGACGATTTTCAGGTTTCTGTCAGAGCAGCGGTGCCAGAAAGAGGT 5076
Qy 1628 LysAsnLysSerGluAspAlaLysSerValHisPheGluSerSerThiSerGluSerSer 1647
Db 5077 AAAAATTAAGTCAGAAAGATGCAAAAGCTGTTCATTTGATCCACACCTCAGAGTCTCTCA 5136
Qy 1648 ValLeuSerSerSerSerProGluSerThiLeuValLysProGluProAsnGlyIleThi 1667
Db 5137 GTGCTATCAAGTATAGTCCAGAGAGTACCTGTTGTAACCAAGAGCCGAATGGCAATACC 5196
Qy 1668 IleProGlyIleSerSerAspValProGluSerAlaHisLysThiThiAlaSerGluAla 1687
Db 5197 ATCCCTGGTATCTCTTCAAGATGTGCCAGAGGTGCCCAAAACCTACGTGCTCAGAGCA 5256
Qy 1688 LysSerAspThiArgLysInProThiLysValGlyArgPheGlnValThiThiThiAlaAsn 1707
Db 5257 AAGTCACAGACTGGGCAAGCTTACCAAGAGTGGAGGTTTTCAGGAGCAACCTACAGCAAA 5316
Qy 1708 LysValGlyArgPheSerValSerLysThiGluAspLysIleThiAspThiLysLysGlu 1727
Db 5317 AAAGTGGTGGTCTTCTCTGATATCAAAAACAGACAGATCACTGACCAAAAGAAAGAA 5376
Qy 1728 GlyProValAlaSerProProPheMetAspLeuGlnAlaValLeuProAlaValIle 1747
Db 5377 GGACAGAGGCACTCTCTCTTATGATTTGGAACAAAGCTGTTCTTCCGTGCTGATATA 5436
Qy 1748 ProLysLysGluLysProGluLeuSerGluProSerHisLeuGlnLysProSerSerAsp 1767
Db 5437 CCAAAAGAAAGAGAGCTGAACTGTACAGAGCTTCACTATTAATGAGCGCTTCTGAC 5496
Qy 1768 ProGluAlaAlaPheLeuSerArgAspValAlaAspGlySerGlySerProHisSerPro 1787
Db 5497 CCGAGAGCGGCTTTTATTAAGGAGATGGATGGATGGATGGATGGATGGATGGATGGATGG 5556
Qy 1788 HisGlnLeuSerSerLysSerLeuProSerGlnLeuLeuSerGlnSerLeuSerAsnSer 1807
Db 5557 CATCAGTACAGCTCAAAAGAGCTTCTTACCCAGAAATCTTAAGTCAAAAGCTTATGTAATTA 5616
Qy 1808 PheAsnSerSerThiYrMetSerSerAspAsnLysSerAspIleGluAspGluAspLeuLys 1827
Db 5617 TTTAATCTCTCTTACATAGATAGGACACATAGATAGATAGATAGATAGATAGATAGATAG 5676
Qy 1828 LeuGluLeuArgArgLeuArgAspLysHisLeuLysGluLysLeuAspLeuGlnSerArg 1847
Db 5677 TTAGAGCTGCAGACATACGAGATTAACATCTCAAAAGATTCAGAGCTGACAGCTGACGTC 5736
Qy 1848 GlnLysHisGluIleGluSerLeuThiThiLysLeuGlyLysValProProAlaValIle 1867
Db 5737 CAGAAAGCATGAATTTGAATCTTTGTATACCAAACTGGGCAAGGTGCCCTGCTGTATTT 5796
Qy 1868 IleProProAlaAlaProLeuSerGlyArgArgArgArgProThiLysSerLysGlySer 1887
Db 5797 ATTCCTCCAGCTGCTCCCTTTCAGGAGAGAGACGACCACTTAAAGCAAAAGCAGCAGC 5856
Qy 1888 LysSerSerArgSerSerSerSerLeuGlyAsnLysSerProGlnLeuSerGlyAsnLeuSer 1907
Db 5857 AAATCTAGTCAGAACAGATCTCTTGGGGAATTAAGCCCTTCAGGTTCAGGTACCTGTCT 5916
Qy 1908 GlyLysSerAlaAlaSerValLeuHisProGlnGlnThiLeuHisProProGlyAsnIle 1927
Db 5917 GGTCAAGATGAGCTTCACTTGTTCACCCCAAGACACCTTCCCTCTGGCAACATC 5976
Qy 1928 ProGluSerGlyGlnAsnGlnLeuLeuGlnProLeuLysProSerProSerSerAspAsn 1947
Db 5977 CCAGAGTCCGGGCAAGATCAGCTGTTCAGCCCTTAAGCCATCTCCCTCAGTACACAC 6036
Qy 1948 LeuThiSerAlaPheThiSerAspGlyAlaAlaIleSerValProSerLeuSerAlaProGly 1967
Db 6037 CTCTATTCAGCTTACCAAGATGTGTCATTTCAATTCACCAACCTTCTGCTCCAGGT 6096
Qy 1968 GlnGlyThiSerSerThiThiThiThiValIleAlaThiValAsnSerGlnAlaAlaGlyAla 1987
Db 6097 CCAAGAACCAAGCAGCAAAACATCTGTTGGGCAACATGAAACCCCAAGCCCAAGCT 6156
Qy 1988 GlnProProAlaMetThiSerSerArgLysGlyThiPheThiAspAspLeuHisLysLeu 2007
Db 6157 CAGCTCTCTCCATGACGTCCAGCAGGAGAGGACATTCACAGATGACTTGCACCAAGTTG 6216
Qy 2008 ValAspAsnThiAlaArgAspAlaMetAsnLeuSerGlyArgArgGlySerLysGlyHis 2027
Db 6217 GTTACCAATTTGGGCGCCAGATGCCATTAATCTTCAAGCAGGAGAGGAAACCAAGGGCAC 6276
Qy 2028 MetAsnThiGluGlyProGlyMetAlaArgLysPheSerAlaProGlyGlnLeuCysIle 2047
Db 6277 ATGATTAATAGAGGCCCTTGAATGGCAAGAAAGTCTCTGCACTGGGCAACTGTGATC 6336
Qy 2048 SerMetThiSerAsnLeuGlyLysSerAlaProIleSerAlaAlaSerAlaThiSerLeu 2067
Db 6337 TCCATGACCTCGAACCCTGGGTGCTGCCCCCATCTCTCAGAGATCAGCTACCTCTCTA 6396
Qy 2068 GlyHisPheThiLysSerMetCysProProGlnGlnThiPheProAlaThiProPhe 2087
Db 6397 GGTACACTTCAACAAAGTATATGTGCCCCCAACAGATATGCTTTCCAGCTACCCCAATTT 6456
Qy 2088 GlyAlaGlnThiPheSerGlyThiGlyGlyProAlaProGlnProLeuGlnLysGlnPro 2107
Db 6457 GGGCTCAATGGAGTGGAGCGGGTGGCCAGACCAAGCACTTGGCCAGTTCACACT 6516
Qy 2108 ValGlyThiAlaSerLeuGlnAsnPheAsnLysSerAsnLeuGlnLysSerIleSerAsn 2127
Db 6517 GTGGGAACCTGCTCTTGGCAAAATTCACATCAGCAATTTGCAAAATCCATCAGCAAC 6576
```


QY	561	ArgValSerLeuLleIysArgIysArgGluInaArgGluLeuValArgGluGluGluInu	580
Db	1681	AGAGTATCTATTATTATTAAGAGAGAAACAGAGAGCAGCGGCAATTGGTACGGGAGGAGCAAGA	1741
QY	581	LysIleValyGluGluGluSerSerLeuIysGluInaIleGluInSerSerAlaSerGln	600
Db	1741	AAAAAAGCAGGAGAGAGAGAGCTCTCAAAACAGCAGGTAGAACATTCAGTCTTCCGAG	1800
QY	601	ThrGlyIleIysGluInleuProSerAlaSerThrGlyIleProThrAlaSerThrThrSer	620
Db	1801	ACAGGAATCAGAGAGCTCCCTCTGTGTAGCACCGGCACTACCTACTCTTCTTACCACTTCA	1860
QY	621	AlaSerValSerThrGluValIleGluProGluInuProGluValAlaAspGlnIleGluInleu	640
Db	1861	GGTTCAGTTCTTACACCACTGAGTAACCTGAAAGAACTGAGGAGCATCAACATCAACACTA	1920
QY	641	GlnIYrGlnGlnProSerIleSerValleuSerAspGlyThrValAspSerGlyGlnGly	660
Db	1921	CAGTACCAACCAACCAAGTATCTGTGTATCTGATGGAGCGGTGACAGTGTACAGGA	1980
QY	661	SerSerValPheThrGluSerArgValSerSerGlnInuThrValSerTyrlYserGln	680
Db	1981	TCCTCTGTCTTACAGAAATCTCGAGTGAAGCACCACCAAGACAGTTCATATGTGTTCCAA	2040
QY	681	HisGluGlnAlaHisSerThrGlyThrValProGlyHisIleProSerThrValGlnAla	700
Db	2041	CATGAACAGGACCAATTTTACAGGACAGAGTCCAGGGCATATATCTTCTTACTGTCCAAACA	2100
QY	701	GlnSerGlnProHisGlyValTyrrProProSerSerValAlaGlnGlyGlnSerGlnIy	720
Db	2101	CAGTCTCAGCCCAATGGGTATATTCACCCCTCAAGTGTG-----	2139
QY	721	GlnProSerSerSerSerleuThrGlyValSerSerSerGlnProIleGlnHisProGln	740
Db	2139	-----	2139
QY	741	GlnGlnGlnGlnIyIleGlnGlnInuThrAlaProProGlnGlnInuThrValGlnTyrrSerleuSer	760
Db	2140	---CAGCAGGGAGATACAGAGACAGACCCCTCTCTCAACAGACAGTGCAGATATTCATTTCA	2196
QY	761	GlnThrSerThrSerSerGlnValaThrThrAlaGlnProValSerGlnProGlnAlaPro	780
Db	2197	CAGATCATCACTCCAGTGAAGGCCACTACTGCACAGCCAGTGAAGTCAAGCTCTCAAGCTCA	2256
QY	781	GlnValleuProGlnValSerAlaGlyIyrsGln-----	791
Db	2257	CAGGTCTTGCCCTCAAGTATCAGCTGGAAAAAGCTTCCAGTTTCCACAGCATCAAACT	2316
QY	791	-----	791
Db	2317	ATCCAAAGCGAAGCTCAAGATCCCACTGTCGACACAAACCTCGTGTGTTCCAGTCCACTCT	2376
QY	791	-----	791
Db	2377	GGTGTCAATTCCCTTCCAGTGGAGACGCGCTCCCTAATCCCTTGTCCTCCAGATACCT	2436
QY	791	-----	791
Db	2437	GTTCTCAGATTCCTCATATCAATCTTCATGTGTCTACGCTCAGACAGAGTTTCTCATCC	2496
QY	791	-----	791
Db	2497	CTTCCATCACAATGGAGCTGGCATTACTAGCTCTGTCTCAGATTGAGCTTCAATCTGCT	2556
QY	791	-----	791
Db	2557	ACAAAGCTGGAGATCCGGGGGTATCAACTGTGGTTCTTAGTCAAGCTTCCAAACCTTTCTG	2616
QY	791	-----	791
Db	2617	CAGCTGTGACTCAGCTGCCAAGTCAGTTCAACCCACAGCTCTTACAAACAGACAGTTCTAG	2676

QY	791	-----	791
Db	2677	TCGATGGAAATACAGACTAACCCTTGGAACAAGCTGAGATTCCACTTTCCTGTGGAGAT	2738
QY	791	-----	791
Db	2737	GTTCTGTACAGAGGCTCCCACTCGACTGCAACCAAGTACCCAGAGATTCAAAATATT	2799
QY	791	-----	791
Db	2797	GCTCCCTCTTCCACGTGGCTTCTGTGTTCATTCATTTCAGAGTCCCTTCCCATG	2855
QY	791	-----	791
Db	2857	CCGACAGAACTGAGCTACACCTGGATCTTCCACAGTGTGCAGCTTATGTGAA	2916
QY	791	-----	791
Db	2917	TCAAAATCTTTAGTTCCTATGGGTGTGTAGAGGACAGTTCAAGTGTCCAGCCAGAA	2978
QY	792	-----	792
Db	2977	GGAAGTTTAGCACAAGCCCCCACTACATCCCTCCGACAAGACAGTTTGGAGAGTCTAG	3038
QY	795	GlyValSerGlnValAlaProAlaGluProValAlaValAlaGlnProGlnAlaThrGln	814
Db	3037	GGAAGTCTCAGAGTGTCTCTGACAGAGCAGTGTGACATGACAGCCCCAAGCTTCCAG	3098
QY	815	ProThrThrLeuAlaSerSerValAspSerAlaHisSerAspValAlaSerGlyMetSer	834
Db	3097	CCGACCACTTTGGCTTCTCTGTACACAGTGACATTGAGATGTGCTTCAGATATGAT	3156
QY	835	AspGlyAsnGluAsnValProSerSerSerGlyArgHisGluGlyArgThrThrLysArg	854
Db	3157	GATGGCAATGAGAACGTCCCATCTTCCAGTGGAAAGCATGAGAGAACTACAAAACGG	3216
QY	855	HisLysArgLysSerValArgSerArgSerArgHisGluLysThrSerArgProLysLeu	874
Db	3217	CATTACCCAAAATCTGTAGAGAGTGTCTGTGACATGAAAACCTTCACGCCAAAATTA	3276
QY	875	ArgLLeuAsnValSerAsnLysGlyAspArgValValGluCysGlnLeuGlnThrHis	894
Db	3277	AGAAATTTGAAATGTTCAATTAAGAGAACCCGAGTAGAGATGCAATTAGAGACTCAT	3338
QY	895	AsnArgLysMetValThrPheLysPheAspLeuAspGlyAspAsnProGluGlnLea	914
Db	3337	AATAGAAAATGTGTACATTCAAATTTGACCTAGATGTGTACAACCCCGAAGAAATAGCA	3398
QY	915	ThrLLeuValAsnAsnAspPheLLeuAlaIleGluArgGluSerPheValAspGln	934
Db	3397	ACAAATATGAGTGAACAATGACTTTATTTTACCAATAGAGAGAGAGTGTGTGTGATCAA	3456
QY	935	ValArgGluIleIleGluLysAlaAspGluMetLeuSerGlyAspValSerValGluPro	954
Db	3457	GTCGAGAAATTTATGAAGAGCTGATGAATGCTCAGTGGAGATGTGAGTGAACA	3516
QY	955	GluGlyAspGlnGlyLeuGlnSerLeuGlnGlyLysAspArgLysPheSerGlySer	974
Db	3517	GAGGCTGATCAAGGATTTGGAAGTCAACAAGAAAGATGACTATGTGCTTTCAGTTCT	3576
QY	975	GlnLysLeuGluGlyGluPheLysGlnProLLeuProAlaSerSerMetProGlnGlnLe	994
Db	3577	CAGAAATTTGAGAGGAGAGTTCMAACAACCAATTCCTGTGCTTCCATGCCCACACAATA	3636
QY	995	GlyLLeuProThrSerSerLeuThrGlnValValHisSerAlaGlyArgArgPheIleVal	1014
Db	3637	GGCATTCCTACAGTCTTTTAATCTCAAGTGTTCATTCCTGGGGAAGCGGTTTATATGTG	3696
QY	1015	SerProValProGluSerArgLeuArgLysSerLysValPheProSerGluLLeuThrAsp	1034
Db	3697	AGTCTGTGCCAGAAAGCCGATTTGACGAAGATCAAAAATTTTCCCATGTGAATAACAGAT	3756
QY	1035	ThrValAlaAlaSerThrAlaGlnSerProGlyMetAsnLeuSerHisSerAlaSerSer	1054

```
Db 3757 ACACTGCTGCTCTACAGCTCAGAGCCCTGGATGAACTGTCTCACTCTGCATCTCC 3816
Qy 1055 LeuSerLeuGlnGlnAlaPheSerGluLeuArgAlaGlnMetThrGlyProAsn 1074
Db 3817 CTGAGCTACAAAGAGGCTTTCTGAACTTAGACCTGCCCCAAATGACAGAGAGCCCAAY 3876
Qy 1075 ThrAlaProProAsnPheserHisThrGlyProThrPheProValProProPheLeu 1094
Db 3877 AAGAGACTCCAAACTTAGTCATACAGACCAACATTCAGATGATGACCTCTTTCTTA 3936
Qy 1095 SerSerIleAlaGlyValProThrThrAlaAlaAlaThrAlaProValProAlaThrSer 1114
Db 3937 AGTAGCATGCTGGAGTCCCAACACACAGAGAGCCACAGACCAAGCTCCGCAAAAGC 3996
Qy 1115 SerProProAsnAspIleSerThrSerValIleGlnSerGluValThrValProThrGlu 1134
Db 3997 AGCCCTCTCAATGACATTCACATCACTCAATTCAGTCTGAGGTACAGTCCCACTGAA 4056
Qy 1135 GluGlyIleAlaGlyValAlaAlaThrSerThrGlyValAlaThrSerGlyLeuProIle 1154
Db 4057 GAGGGGATGCTGGAGTCCACACAGACAGGTGTGTAATCTTCAAGTGTCTCCCATTA 4116
Qy 1155 ProProValSerGluSerProValLeuSerSerValValSerSerIleThrIleProAla 1174
Db 4117 CCACCTGTCTGATCAACAGTACTTTCAGCCGTAGTTTCAAGTATCAACATACCTGCA 4176
Qy 1175 ValValSerIleSerThrThrSerProSerIleGlnValProThrSerThrSerGluIle 1194
Db 4177 GTTGCTCAATATCACTACATACATCCCTGACCTCAAGTCCCATCCATCCGAGATC 4236
Qy 1195 ValValSerSerThrAlaLeuThrProSerValThrValSerAlaThrSerAlaSerAla 1214
Db 4237 GTTGCTTCTAGTACAGCATGTATCTTCACTACAGATTCAGCACTTCCAGCTCTGCA 4296
Qy 1215 GlyGlySerThrAlaThrProGlyProIysProProAlaValAlaSerGlnGlnAlaAla 1234
Db 4297 GGGGGCACTAGTACTACCCAGAGTCTTAAGCTCCAGCTGATGATTCAGAGAGGACCA 4356
Qy 1235 GlySerThrThrValGlyAlaThrLeuThrSerValSerThrThrThrSerPheProSer 1254
Db 4357 GGCAGCACTACTGTGGAGCCACATTAACATCACTTACACACATCTTCCCAAGC 4416
Qy 1255 ThrAlaSerGlnLeuSerIleGlnLeuSerSerThrSerThrProThrIleuAlaGln 1274
Db 4417 ACACTTACAGCTCTCACTTACGCTTACAGCACTTCTTACTCTTACTTACTTAA 4476
Qy 1275 ThrValValAlaSerAlaHisSerLeuAspIysThrSerHisSerSerThrThrGlyLeu 1294
Db 4477 ACCGTGATGTTAGGACACACTCACTAGTAAAGATCATCTCATAGCAACTGATG 4536
Qy 1295 AlaPheSerLeuSerAlaProSerSerSerSerSerProGlyAlaGlyValSerSerIyr 1314
Db 4537 GCTTCTCTCTCTGACCACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4596
Qy 1315 IleSerGlnProGlyGlyLeuHisProLeuValIleProSerValIleAlaSerThrPro 1334
Db 4597 ATTTCTACGCTGTGGGCTGTATCTTTGGTCACTTCCATCAGTATAGCTTCTACTCT 4656
Qy 1335 IleLeuProGlnAlaAlaGlyProThrSerThrProLeuLeuProGlnValProSerIle 1354
Db 4657 ATTCTCCCAAGAGAGAGCACTTCTTACACCTTTATATACCCCAAGTACTGATATC 4716
Qy 1355 ProProLeuValGlnProValAlaAsnValProAlaValGlnGlnThrLeuIleHisSer 1374
Db 4717 CCACCTTGTGTAACACTGTGTCATATGCTGTGTAACAGCAACACTAATTTATAGT 4776
Qy 1375 GlnProGlnProAlaLeuLeuProAsnGlnProHisSerHisSerProGlyValAlaAspSer 1394
Db 4777 CAGCTCAACAGCTTTGCTTCCCAACAGCCCACTACTCATTTGCTCTGAGTAACTTCT 4836
Qy 1395 AspThrGlnProIysAlaProGlyIleAspAspIleIysThrLeuGluGluIysLeuArg 1414
Db 4837 GATACAGACCCCAAGCTCTGGAATTGATGACATAAAGACTTCAAGAAAGAAAGCTGGC 4896
Qy 1415 SerLeuPheSerGluHisSerSerSerGlyAlaGlnHisAlaSerValSerLeuGluThr 1434
Db 4897 TCTCTGTTAGTGAACACAGCTCATCTGAGCTCAGACATGCTGTCTCATCTGAGACC 4956
Qy 1435 SerLeuValIleGlnSerThrValThrProGlyIleProThrThrAlaValAlaProSer 1454
Db 4957 TCACTAGTCATAGAGACACTGTACACAGAGCACTCCCACTTACTGCTGTGACACAGC 5016
Qy 1455 LysLeuLeuThrSerThrThrSerThrCysLeuProProThrAsnLeuProLeuGlyThr 1474
Db 5017 AAACCTCGACTTCTTACACACAGATCTGTACACCAACCAATTAACCTAGGAACA 5076
Qy 1475 ValAlaLeuProValIleThrProValValThrProGlyGlnValSerThrProValSerThr 1494
Db 5077 GTTGCTTGGCAGTTACACCAAGGTGTCACCTGGCAAGTTCTTACCCAGTACGACT 5136
Qy 1495 ThrThrSerGlyValLysProGlyThrAlaProSerLysProProLeuThrIleAlaPro 1514
Db 5137 ACTACATCAGAGAGTGAACCTGGAAGCTCTCCCAAGCCACCTCTAATAGGCTCG 5196
Qy 1515 ValLeuProValGlyThrGluLeuProAlaGlyThrLeuProSerGluGlnLeuProPro 1534
Db 5197 GTGCTGCAGAGGTGACTGAATTCAGAGATCTTCAACCTCAGCAGCAGCTGCACT 5256
Qy 1535 PheProGlyProSerLeuThrGlnSerGlnProLeuGlnAspLeuAspAlaGlnLeu 1554
Db 5257 TTTCAGGACTTCTTACACCAAGTCCAGCACTCTTACAGAGATCTTGAATGCTCAATG 5316
Qy 1555 ArgArgThrLeuSerProGly**IleThrValThrSerAlaValGlyProValSerMet 1574
Db 5317 AGAAGAACACTTATCCAGAGATKATACAGTACTTCTGCGTGTGCTGTCATG 5376
Qy 1575 AlaAlaProThrAlaIleThrGluAlaGlyThrGlnProGlnIleGlyValSerGlnVal 1594
Db 5377 GCGGCTCCAAAGCAATCCAGAAAGAGCAAGCAAGCTCAGAAAGGGTTCCTCAATG 5436
Qy 1595 LysGluGlyProValLeuAlaThrSerSerGlyAlaGlyValPheLysMetGlyArgPhe 1614
Db 5437 AAGAGAGCCTGTCTTACAGCACTAGTTCAGAGAGTGTGTCTTTTACATGGAGCAATT 5496
Qy 1615 GlnValSerValAlaAlaAspGlyAlaGlnLysGluGlyLysAsnLysSerGlnAspAla 1634
Db 5497 CAGCTTCTGTGACAGAGAGGTCGCCAAGAGGGTAAATAATCACTAGAGATCA 5556
Qy 1635 LysSerValHisPheGluSerSerThrSerGluSerSerValLeuSerSerSerSerPro 1654
Db 5557 AAGTCTGTTCATTTGAAATCCAGACCTCAGAGTCTTCAAGTCAATGATGTC 5616
Qy 1655 GluSerThrLeuValLysProGlnProAsnGlyIleThrIleProGlyIleSerSerAsp 1674
Db 5617 GAGGATACCTGTGGTAAACAGAGCCGAAATGGCATACCACTCCGTGATCTTCAAT 5676
Qy 1675 ValProGluSerAlaHisLysThrThrAlaSerGlnAlaLysSerAspThrGlyGlnPro 1694
Db 5677 GTGCAGAGAGTGCACCAAAACTACGCTCAGAGGCAAGTCAAGACATGGGAGCCT 5736
Qy 1695 ThrLysValGlyArgPheGlnValThrThrThrAlaAsnLysValGlyArgPheSerVal 1714
Db 5737 ACCAAGTGTGACGTTTCAAGTGAACACTACAGAAACAAAGTGTGCTTCTGTGA 5796
Qy 1715 SerLysThrGluAspLysIleThrAspThrLysLysGluGlyProValAlaSerProPro 1734
Db 5797 TCAAAAACCTGAGACAGATCACTGACCAAGAAAGAAAGACAGTGGCATCTCTCT 5856
Qy 1735 PheMetAspLeuGlnGlnAlaValLeuProAlaValIleProLysLysGluLysProGln 1754
Db 5857 TTATGATTTGGAACAGAGCTTCTCTCTGTGTGATACCAAGAAAGAGAGAGCTGAA 5916
Qy 1755 LeuSerGluProSerHisLeuAsnGlyProSerSerAspProGluAlaAlaPheLeuSer 1774
Db 5917 CTGTCAAGCCTTCACTAATATGCGCTCTTCTGACCGGAGCGCGCTTTTAACT 5976
```


QY 261 GluArgGlnArgPheLysGluGluAlaGluMetLeuLysGlyLeuGlnHisProAsnIle 280
DB 601 GAGAGGACAGAGATTAAAGAAAGCTGAATGTTAAAGGTCTTCAGACATCCCAATATT 660
QY 281 ValArgPheTyrAspSerTyrGluSerThrValLysGlyLysLysCysIleValLeuVal 300
DB 661 GTTAGATTATTAATGATTCCTGGGAATCCACAGTAAAGAAAGAGTGCATGTTTGGTG 720
QY 301 ThrGluMetThrSerGlyThrLeuLysThrTyrLeuLysArgPheLysValMetLys 320
DB 721 ACTGAACCTTATGACGTCTGGAACCTTAAACGTATCTGAAGAGTTTAAAGTATGTAAG 780
QY 321 IleLysValLeuArgSerTyrCysArgGlnIleLeuLysGlyLeuGlnPheLeuHisThr 340
DB 781 ATCAAGATCTTAAGAGCTGGTGGCGTCAAGTCTTAAAGGCTTCAGATTCTTCATCT 840
QY 341 ArgThrProProIleIleHisArgAspLeuLysCysAspAsnIlePheIleThrGlyPro 360
DB 841 CGAATCCACCTATCATTCACCGGAGTCTTAAATGTGACCAATCTTATCACCGGCTT 900
QY 361 ThrGlySerValLysIleGlyAspLeuGlyLeuAlaThrLeuLysArgAlaSerPheAla 380
DB 901 ACTGCTCAGTCAAGATTGAGACCTCGGTCTGGCAACCTGAAGCGGGCTTCTTTTGGC 960
QY 381 LysSerValIleGlyThrProGluPheMetAlaProGluMetTyrGluGluLysTyrAsp 400
DB 961 AAGAGTGTATAGTATCCCGACAGATTCATGGCCCTGTGAAGTATGAGAGAAATATAT 1020
QY 401 GluSerValAspValTyrAlaPheGlyMetCysMetLeuGluMetAlaThrSerGluTyr 420
DB 1021 GAATCCGTGTAAGTTATGCTTTTGGATGTGCATGCTTGAATGGTCAATCTGCAATAT 1080
QY 421 ProTyrSerGluCysGlnAsnAlaAlaGlnIleTyrArgArgValThrSerGlyValLys 440
DB 1081 CCTTACTCGAGTGGCAAAATGCTGCACAGATCTACCGTGGGTGACCAAGGGGTGAG 1140
QY 441 ProAlaSerPheAspLysValAlaIleProGluValLysGluIleIleGluGlyCysIle 460
DB 1141 CCAAGCACTTGGACAAAGTACCAATCTCTGATGAGAAAGAAATTAATGAAAGATGCA 1200
QY 461 ArgGlnAsnLysAspGluArgTyrSerIleLysAspLeuLeuAsnHisAlaPhePheGln 480
DB 1201 CGACAAAACAAGATGAAGAATATTCATCAAGACCTTTTGAACCATGCTTCTTCCA 1260
QY 481 GluGluThrGlyValArgValGluLeuAlaGluGluAspGlyGluLysIleAlaIle 500
DB 1261 GAGGAAACAGAGATCGGGTAGAATTAGCAGAAAGATGATGAGAAATAATACCATTA 1320
QY 501 LysLeuTyrPheAspArgIleGluAspIleLysLysLeuLysGlyLysTyrLysAspAsnGlu 520
DB 1321 AATATTGCTGCTAGATTGAAAGATTAATTAAGAAATTAAGGAAATTAACAAAGATTAATGA 1380
QY 521 AlaIleGluPheSerPheAspLeuGluArgAspValProGluAspValAlaGlnIleMet 540
DB 1381 GGTATTGAGTTTCTTTGATTGAGAGAGATGTCCCAAGAAAGTGTGCAACAAGAAATG 1440
QY 541 ValGluSerGlyTyrValCysGluGlyAspHisLysThrMetAlaLysAlaIleLysAsp 560
DB 1441 GTAGAGCTGGGTATGCTGTGTAAGTGTATCAACAAGCATGTGCTAAAGCTTAAAGC 1500
QY 561 ArgValSerLeuIleLysArgLysArgGluGlnArgGlnLeuValArgGluGluGlnIle 580
DB 1501 AGAGTATCTTAATTAAAGGAAACAGAGCGACCGGAGTTGTATCGGAGAGAGCAAGAA 1560
QY 581 LysLysLysGlnGluGluSerSerLeuLysGlnGlnValGluGlnSerSerAlaSerGln 600
DB 1561 AAAAAAAGCAGAGAGAGAGAGCTCAACAAGCAGAGTGAACAATCCAGTCTTCCAG 1620
QY 601 ThrGlyIleLysGlnLeuProSerAlaSerThrGlyIleProThrAlaSerThrThrSer 620
DB 1621 ACAGAAATCAAGACAGCTCTCTCTGTGTAGACACCGGATCACTACTGCTTCAACACTTCA 1680

QY 621 AlaSerValSerThrGlnValGluProGluGluIleAspGlnHisGlnGlnLeu 640
DB 1681 GCTTCAGTTTCTTACCAAGTGAACCTTGAGAACTTGAAGAGATCAACATCAACACTA 1740
QY 641 GlnTyrGlnGlnProSerIleSerValLeuSerAspGlyThrValAspSerGlyGlnGly 660
DB 1741 CAGTACCAAGCAACCAATATATCTGTGTATCTGATGGACGGTGTGACAGTGTGAGGA 1800
QY 661 SerSerValPheThrGluSerArgValSerSerGlnGlnThrValSerTyrGlySerGln 680
DB 1801 TCTCTGTCTTCAAGAAATCTCGATGAGCAAGCAAGCAAGTTCATATATGTTCCCA 1860
QY 681 HisGluGlnAlaHisSerThrGlyThrValProGluHisIleProSerThrValGlnAla 700
DB 1861 CATGAACAGCAACTTCTACAGCAGCAGCTCCAGGGCATATACCTTCACTGTCACAGCA 1920
QY 701 GlnSerGlnProHisGlyValTyrProProSerSerValAlaGlnGlnGlnSerGlnGly 720
DB 1921 CAGTCTCAGCCCCAGGGGTATATCCACCTCAAGTGTGGCACAGGGGACAGGCCAGGCT 1980
QY 721 GlnProSerSerSerSerLeuThrGlyValSerSerSerGlnProIleGlnHisProGln 740
DB 1981 CAGCATCTCAAGATGCTTAAAGGGGTTCATCTTCCCAACCACTCAACATCTCTCG 2040
QY 741 GlnGlnGlnGlyIleGlnGlnThrAlaProProGlnGlnThrValGlnTyrSerLeuSer 760
DB 2041 CAGCAGCAGGGAATACAGCAGACAGCCCTCTCAAGACAGTGCAGATTAATCACTTCA 2100
QY 761 GlnThrSerThrSerSerGluAlaThrThrAlaGlnProValSerGlnProGlnAlaPro 780
DB 2101 CAGACATCAACCTCAGTGAAGCCCACTACAGCAGCAGAGTGCAGCTCAAGCTTCA 2160
QY 781 GlnValLeuProGlnValSerAlaGlyLysGlnSerThrGlnGlyValSerGlnValAla 800
DB 2161 CAGTCTTGTCTCAAGTATCACTGAGAAACAGAGTATCTCAGGAGTCTCTCAGTTCT 2220
QY 801 ProAlaGluProValAlaValAlaGlnProGlnAlaThrGlnProThrThrLeuAlaSer 820
DB 2221 CCTGAGAGCCAGTTGAGTGAAGCAGCAGCCCAAGTACCCAGCCAGCACTTGGCTTCC 2280
QY 821 SerValAspSerAlaHisSerAspValAlaSerGlyMetSerAspGlyAsnGluAsnVal 840
DB 2281 TCTGTAGACAGTGCACATTCAGATGTTCTCAGGTATGAGTGAAGTGAAGAACGTG 2340
QY 841 ProSerSerSerGlyArgHisGluGlyValArgThrThrLysArgHisTyrArgLysSerVal 860
DB 2341 CCATCTTCAAGTGAAGGACATGAAGAGAACTACAAACCGGCATTACCGAAATCTGTA 2400
QY 861 ArgSerArgSerArgHisGluLysThrSerArgProLysLeuArgIleLeuAsnValSer 880
DB 2401 AAGAGTGGCTCTCGACATGAAGAAACCTTCAAGCCCAAAATTAAGAAATTTGATGTTCA 2460
QY 881 AsnLysGlyAspArgValValGluCysGlnLeuGluThrHisAsnArgLysMetValThr 900
DB 2461 AATAAAGAGACCGAGTGTGAATGTCAATTAAGACTCAATAAAGAAATGTGTTCA 2520
QY 901 PheLysPheAspLeuAspGlyAspAsnProGluGluIleAlaThrIleMetValAsnAsn 920
DB 2521 TTCAATTTGACCTGATGAGTGAACACCCGAGAGATGACATTAATGTGTAACAAT 2580
QY 921 AspPheIleLeuAlaIleGluArgGluSerPheValAspGlnValArgGluIleIleGlu 940
DB 2581 GACTTTATTTAGCAATAGAGAGAGATGCTTGTGTGATCAAGTCCGAGAAATTTTGA 2640
QY 941 LysAlaAspGluMetLeuSerGluAspValSerValGluProGluGlyAspGlnGlyLeu 960
DB 2641 AAAGCTGATGAAGTCTAGTGAAGATGTCAAGTGTGAACCAAGAGGATGATCGAGGATTTG 2700
QY 961 GluSerLeuGlnGlyLysAspAspTyrGlyPheSerGlySerGlnLysLeuGluGluGln 980
DB 2701 GAGAGTCTACAGAGAAAGATCACTATGGCTTTTGAAGTCTCAAGAAATTTGAAGAGAG 2760
QY 981 PheLysGlnProIleProAlaSerSerMetProGlnGlnIleGlyIleProThrSerSer 1000

```
Db      2761 TTCAAAACCAATTCCTCCCTTCATGCCACGAATAGGCAATTCCTACAGTTCT 2820
Qy      1001 LeuThrGlnValValHisSerAlaGlyArgPheIleValSerProValProGlnSer 1020
Db      2821 TTAAGTCAAGTTGTTCTTCCTGGGGAAGCGGTTATAGTACGTCCTGGCCAGAAAC 2880
Qy      1021 ArgLeuArgGlnSerLysValPheProSerGluIleThrAspThrValAlaIleSerThr 1040
Db      2881 CCAATTAACGAGATCAAAAGTTTCCCAAGTAAATACAGATACAGTTGCTGCTCTACA 2940
Qy      1041 AlaGlnSerProGlyMetAsnLeuSerHisSerAlaSerSerLeuSerLeuGlnGlnAla 1060
Db      2941 GCTCAGAGCCCTGGAGATGAACTGTCTCACTCGATCATACCTTAACTTACAAACAGGCC 3000
Qy      1061 PheSerGlnLeuArgAlaGlnMetThrGlnGlyProAsnThrAlaProProAsnPhe 1080
Db      3001 TTTTGTGAACCTTAGAGCTGCCAAATGACAGAGAACCCAAATACAGCACTTCGAAATTT 3060
Qy      1081 SerHisThrGlyProThrPheProValValProPheLeuSerSerIleAlaGlyVal 1100
Db      3061 AGTCATACAGAGCAACCAATTCCTCAGTAGTACTCTTCTTAAGTAGATGCTGGAGTC 3120
Qy      1101 ProThrThrAlaAlaAlaThrAlaProValProAlaThrSerSerProProAsnAspIle 1120
Db      3121 CCAACCAAGAGAGCCACAGCAACAGCAGCCTCTGCAACAAAGCAAGCCCTCTAATGACATT 3180
Qy      1121 SerThrSerValIleGlnSerGlnValThrValProThrGlnGlnGlyIleAlaGlyVal 1140
Db      3181 TCCACATCAGTAAATTCAGTCTGAGGTTACAGTCCCACTGAAGAGGAGATTCCTGAGATT 3240
Qy      1141 AlaThrSerThrGlyValValValThrSerGlyLeuProIleProProValSerGlnSer 1160
Db      3241 GCCACCAAGCAGCAGGAGTGTAATCTTCAGTGGTCTCCCAATACCACTGTGTCTGAATCA 3300
Qy      1161 ProValLeuSerSerValValSerSerIleThrIleProAlaValValSerIleSerThr 1180
Db      3301 CCAGTACTTCCAGGAGTGTTCAGATCAACAAATACCTGCAAGTGTCTCAATATCTACT 3360
Qy      1181 ThrSerProSerLeuGlnValProThrSerThrSerGluIleValValSerSerThrAla 1200
Db      3361 ACATCCCGCTCACTCAAGTCCCAATCCCAATCCCAATCTGAGATGCTGTTCCTAGTACACA 3420
Qy      1201 LeuTyrProSerValThrValSerAlaThrSerAlaSerAlaGlyIleSerThrAlaThr 1220
Db      3421 CTGTAATCCTCAGTAAAGTTTCAAGCACTTCAGACTTCGCAAGGGGAGTACTGTACCC 3480
Qy      1221 ProGlyProLysProProAlaValValSerGlnGlnAlaAlaGlySerThrValGly 1240
Db      3481 CCAGGCTCTAAGCTCCAGCTGTAGTATCTCAGCAGCAGCAGCAGCAGCACTACTGTGGGA 3540
Qy      1241 AlaThrLeuThrSerValSerThrThrThrSerPheProSerThrAlaSerGlnLeuSer 1260
Db      3541 GCCACATTAAATCAATGATTCTTCAACCACTTCATTCCTCCAGAGCAAGCTTCCAGGTGCC 3600
Qy      1261 IleGlnLeuSerSerSerThrThrThrProThrLeuAlaGlnThrValValSerAla 1280
Db      3601 ATTCAAGCTTACAGAGTACTTCTACTCTTACTTAAAGTGAACCGTGATGTAGCCCA 3660
Qy      1281 HisSerLeuAspLysThrSerHisSerSerThrThrGlyLeuAlaPheSerLeuSerAla 1300
Db      3661 CACTCACTAAGATGAAGATCTCATAGCAGTACAACTGATGATGGCTTCTCCCTCTCTGCA 3720
Qy      1301 ProSerSerSerSerSerProGlyValAlaGlyValSerSerTyrIleSerGlnProGly 1320
Db      3721 CCAATCTTCTCTTCTCTCTCTGAGAGAGAGATGCTATTAATTTCTCAACCTGGAGTGG 3780
Qy      1321 LeuHisProLeuValIleProSerValIleAlaSerThrProIleLeuProGlnAlaAla 1340
Db      3781 CTGATCTCTTGTGATCATTCATCAGTATGATCTTACTCTTATCTTCCCAAGCAGCA 3840
Qy      1341 GlyProThrSerThrProLeuLeuProGlnValProSerIleProProLeuValGlnPro 1360
Db      3841 GGACCTACTTCAACCTTATTTACCCCAAGTACTAGTATCCACCTTGGTAGAGCT 3900
Qy      1361 ValAlaAsnValProAlaValGlnIleThrLeuIleHisSerGlnProGlnProAlaLeu 1380
Db      3901 GTTGCCAAATGATGCTGCTGTACAGCAGACAACTTAATCAATGACCTTCAACCTGTTG 3960
Qy      1381 LeuProAsnGlnProHisIleThrHisCysProGlnValAlaAspSerAspThrGlnProLysAla 1400
Db      3961 CTGCCAACAAGCCCACTACTACTTGTCTGAGATGATTCGATACACAACTCAAGCTT 4020
Qy      1401 ProGlyIleAspAspIleLysThrLeuGlnGlnLysLeuArgSerLeuPheSerGlnHis 1420
Db      4021 CCTGAATGTGATGATGAATGAAGCTCTGAGAAAGAAAGTGGGCTCTGTTCAGTGAAC 4080
Qy      1421 SerSerSerGlyValAlaGlnHisAlaSerValSerLeuGlnThrSerLeuValIleGlnSer 1440
Db      4081 AGCTATCTGAGGCTCAGACAGCTCTGTCTCTCACTGAGAGACTCACTAGTCAATGAAGC 4140
Qy      1441 ThrValThrProGlyIleProThrThrAlaValAlaProSerLysLeuLeuThrSerThr 1460
Db      4141 ACTGTCACACAGGATCTCCAACTACTGCTGTGGACCAAGCAAACTCTGACTTCTTACA 4200
Qy      1461 ThrSerThrCysLeuProProThrAsnLeuProLeuGlyThrValAlaLeuProValThr 1480
Db      4201 ACAAGTACTTCTTACCAACCAATTTACACATGAGAAAGTGTCTTGGCAAGTTACA 4260
Qy      1481 ProValAlaThrProGlyGlnValSerThrProValSerThrThrThrSerGlyValLys 1500
Db      4261 CCAAGTGTCAACCTGGGCAAGTTCATCCCAAGTCAACACTACTACATCAGAGATGAAA 4320
Qy      1501 ProGlyThrAlaProSerLysProProLeuThrIleAlaProValLeuProValGlyThr 1520
Db      4321 CTTGGAACTGCTCCCTCCCAAGCAACCTTAACTAAGCTCCGGTCTGCCAGTGGTACT 4380
Qy      1521 GluLeuProAlaGlyThrLeuProSerGlnLeuProProPheProGlyProSerLeu 1540
Db      4381 GAACCTTCAGCAGGATCTTACCCAGCAGCAGCTGCCACCTTTCAGAGACTTCTCTA 4440
Qy      1541 ThrGlnSerGlnGlnProLeuGlnAspLeuAspAlaGlnLeuArgThrLeuSerPro 1560
Db      4441 ACCAGTCCCAAGCAACCTTGAAGATCTTATGCTCAATGAGAAAGAACCTTAGTCCA 4500
Qy      1561 Glu**1IleThrValThrSerAlaValGlyProValSerMetAlaAlaProThrAlaIle 1580
Db      4501 GAGATKATCAAGAGACTTTCGGGTTGTCTGTGTCCAGGGGCTCCCAACAGCAATC 4560
Qy      1581 ThrGlnAlaGlyThrGlnProGlnLysGlyValSerGlnValLysGlnGlyProValLeu 1600
Db      4561 ACAGAGCAGGAAACAGACCTCAGAAAGGTGTTTCTCAAGTCAAAAGAGCCCTGCTTA 4620
Qy      1601 AlaThrSerSerGlyAlaGlyValPheLysMetGlyArgPheGlnValSerValAlaAla 1620
Db      4621 GCAACTGATGAGAGCTGGGTTTAAAGATGGAGCAATTCAGGTTTCTGTGACAGA 4680
Qy      1621 AspGlyAlaGlnLysGlnGlyLysAsnLysSerGlnAspAlaLysSerValHisPheGln 1640
Db      4681 GAGGCTGCCAGAAAGAGGGGATTAAGTCAAAATGCAAAATGCAAACTGTTCAATTTGAA 4740
Qy      1641 SerSerThrSerGlnSerSerValLeuSerSerSerSerProGlnSerThrLeuValLys 1660
Db      4741 TCCAGCACTTCAGAGTCTCTCAGTCTCATCAAGTGTAGTCCAGAGAGTACTTGTTGAAA 4800
Qy      1661 ProGlnProAsnGlyIleThrIleProGlyIleSerSerAspValProGlnSerAlaHis 1680
Db      4801 CCAGAGCCGAATGCAATACCATCTCTGTATCTCTTCAAGTGTGCCAGAGAGTGCAC 4860
Qy      1681 LysThrThrAlaSerGlnAlaLysSerAspThrGlnProThrLysValGlyArgPhe 1700
Db      4861 AAAACTGCTGCTCAGAGGCAAAAGTCAAGACTGGGAGGCTTACCAAGGTTGACGTTT 4920
Qy      1701 GlnValThrThrThrAlaAsnLysValGlyArgPheSerValSerLysThrGlnAspLys 1720
Db      4921 CAGGTGACAACTACAGCAAAAGTGGTGTCTTCTGTATCAAAAACCTGAGAGCAAG 4980
```



```

QY 221 LeuYsPheAspIleGluIleGlyArgGlySerPheIleThrValTyrIleGlyLeuasp 240
Db 461 CTCAAGTTGACATGAAATCGGACAGAGCTCTTTAAAGCGGTCTACAAAGGTCTGAC 540
QY 241 ThrGluThrValGluValAlaIleProCysGluLeuGlnAspArgIleLeuThrIleSer 260
Db 541 ACTGAAACCCACCGTGAAGTCCCTGGTGTGAATCGAGGATCGAAATTTAACTAAAGTCT 600
QY 261 GluArgIleArgPheIleGlyGluIleGluIleGluIleGluIleGluIleGluIle 280
Db 601 GAGAGCGACAGATTTAAAGAGAGAGCTGAAATCTTTAAAGGCTTCACAGATCCCAATTT 660
QY 281 ValArgPheTyrAspSerTyrGluSerThrValIleGlyIleGlyIleValIleVal 300
Db 661 GTTAATATTTATGATTTCTCGGAAATCCACAGTAAAGAGAGATGATTTGTTGGTG 720
QY 301 ThrGluLeuMetThrSerGlyThrIleuIleuIleuIleuIleuIleuIleuIleuIleu 320
Db 721 ACTGAATTTATGACGTCTGAAACACTTAAAGATGAAAGGTTTAAAGTGAATGAAG 780
QY 321 IleIleValIleuArgSerTyrCysArgGlnIleLeuIleGlyLeuGluPheLeuIleThr 340
Db 781 ATCAAGTTCTAAGAGAGCTGGGCGGTCAAGATCTTTAAAGGCTTCATCTTCTATCT 840
QY 341 ArgThrProProIleIleHisArgAspLeuIleCysAspAsnIlePheIleThrGlyPro 360
Db 841 CGAATCTCACCTATCATTCACCGCGATCTTAATGTGAACAATCTTTATCATCCGCGCT 900
QY 361 ThrGlySerValIleGlyAspLeuIleGlyLeuIleuIleuIleuIleuIleuIleuIleu 380
Db 901 ACTGCTAGTCAAGATGAGAGATCGGCTCGGAGACCTGAAAGCGGCTCTTTTTCCT 960
QY 381 LysSerValIleGlyThrProGluPheMetAlaProGluMetTyrGluIleuIleuIleuIleu 400
Db 961 AAGAGTGTGATAGTATCCCAAGAGTTCATGCGCCCTGAGATGTATAGAGAGAAATGAT 1020
QY 401 GluSerValAspValTyrAlaPheGlyMetCysMetLeuGluMetAlaThrSerGluTyr 420
Db 1021 GAATCCGTTGACGTTTATGCTTTGGATGTGATGCTTGAAGATGCTGATCATCTGAATAT 1080
QY 421 ProTyrSerGluCysGlnAsnAlaIleGlnIleTyrArgValIleThrSerGlyValIle 440
Db 1081 CCTTACTCGAGTGCAGAAATGCTGCACAGATCTTACCGTCCGAGACAGAGGCGGTGAG 1140
QY 441 ProIleSerPheAspIleValAlaIleProGluValIleGlyIleGluIleGlyCysIle 460
Db 1141 CCAGCCAGTTTGAACAAGTGAAGATTCCTGATGAGAGAGAAATTTATGAAAGATGCA 1200
QY 461 ArgGlnAsnIleAspGluArgTyrSerIleLysAspLeuLeuAsnIleAlaPhePheGln 480
Db 1201 CGACAAAACAAAGATGAAAGATATTCATCAAGACCTTTGAACCATGCTTCTTCCA 1260
QY 481 GluGluThrArgValArgValGluLeuAlaGluIleuAspArgIleGluIleAlaIle 500
Db 1261 GAGGAAACAGAGATGAGGATGAAATGACAGAAAGATGATGAGAAAGAAATATGCCATA 1320
QY 501 LysLeuThrPheuArgIleGluAspIleLysLysLeuIleGlyIleTyrLysAspAsnGlu 520
Db 1321 AAATATATGCTACGTATTTGAAGATTTAAGAAATTTAAAGGAAATATCAAGATTAATGA 1380
QY 521 AlaIleGluPheSerPheAspLeuGluArgAspValProGluAspValAlaGluIleuMet 540
Db 1381 GCTATTAAGTTTCTTTGATTTAGAGAGATGTCCAGAAAGATGTTGACCAAGAAATG 1440
QY 541 ValGluSerGlyTyrValCysGluIleAspIleSerThrMetAlaIleAlaIleLysAsp 560
Db 1441 GTAGAGTCTGGATGCTGTGAAGGTGATCAAGAACCATGGCTTAAAGCTATCAAGAC 1500
QY 561 ArgValSerLeuIleLysArgIleArgGluIleArgGluLeuValArgGluGluGlu 580
Db 1501 AGACTATCATTAATTAAGAGAAACGAGACGCGAGTTGGTACGGAGAGAGCAAGAA 1560

```

```

QY 561 LysIleGluGluGluSerSerLeuIleGluIleGluIleGluIleGluIleGluIleGluIle 600
Db 1561 AAAAAAGCAGAGAGAGAGAGTCTCAACAGCAGGTAGAGAAATCCAGTCTTCCAG 1620
QY 601 ThrGlyIleGluGluProSerAlaSerThrGlyIleProThrAlaSerThrThrSer 620
Db 1621 ACAGGATCAAGAGAGTCTCTCTGCTGATGACACCGGACATACCTACTGCTTACCACTTCA 1680
QY 621 AlaSerValSerThrGluIleGluProGluIleProGluIleAspGluIleGluIleu 640
Db 1681 GCTTCAGTTTCTTCAACAATGAACTGAAAGACCTGAGGACGATCAACATCAACACTA 1740
QY 641 GlnTyrGlnGlnProSerIleSerValLeuSerAspGlyThrValAspSerGlyGlnGly 660
Db 1741 CAGTACCAAGCAACCAAGATATCTGTATATCTATGAGAGCGTTGAGAGGTGACAGGA 1800
QY 661 SerSerValPheThrGluSerArgValSerSerGlnThrValSerTyrGlySerGln 680
Db 1801 TCCCTGCTTCAAGAGATCTCGAGTGCAGCCAAACAGACAGTTTCATATGTTCCCA 1860
QY 681 HisGluGluAlaHisSerThrGlyThrValProGlyHisIleProSerThrValGluAla 700
Db 1861 CATGACGACCATTTCTACAGGACAGTCCAGGCAATATACCTTACTGTTCCAAACA 1920
QY 701 GluSerGlnProIleGlyValTyrProProSerSerValAlaGlnGlyGlnSerGlnGly 720
Db 1921 CAGTCTAGCCCATGGGTATATCCACCTCAAGTGTGACAGAGGCGACAGCCAGGT 1980
QY 721 GlnProSerSerSerSerLeuThrGlyValSerSerSerGlnProIleGlnHisProGln 740
Db 1981 CAGCATCTTCAGATGATTAACAGAGGTTTCATCTTCCCAACCATCAACATCTCAG 2040
QY 741 GlnGlnGlnGlyIleGlnGlnThrAlaProProGlnGlnThrValGlnTyrSerLeuSer 760
Db 2041 CAGCAGAGGAAATACAGCAGACAGCCCTCTCAAGACAGAGTATTCATCTTCA 2100
QY 761 GlnThrSerThrSerSerGluAlaThrThrAlaGlnProValSerGlnProGluAlaPro 780
Db 2101 CAGACATCAACCTCCAGTGAAGGCCACTATGCAACAGCAGATGACCTCAAGCTCCA 2160
QY 781 GlnValLeuProGlnValSerAlaGlyLysGln----- 791
Db 2161 CAAGTCTTGCTCAAGATACAGCTGGAAGAGGCTTCCACCTCGACTGCCACACAG 2220
QY 791 ----- 791
Db 2221 TACCCAGAGATTAATATATTCCTCTCTTCCACAGTGGCTTGTGATTCATCT 2280
QY 791 ----- 791
Db 2281 ACAGTCTTAMCCCTCCCATGCGCAGAGAGTACTGAGTACACTGGGTACTTCCACA 2340
QY 791 ----- 791
Db 2341 GTGGTGAGCCTTATGTGAATCAAAATCTTTAGTCTTATGGTGGTGTAGAGAGACAG 2400
QY 791 ----- 791
Db 2401 GTTCAAGTGTCCAGCCAGAGAGGAGATTGACACAGAGACCCCACTACACTCTCCAGCAA 2460
QY 792 -----SerThrGlnGlyValSerGlnValAlaProAlaGluProValAlaVal 807
Db 2461 GCAGTTTGGAGAGTACTCAGGAGGTCTTCAGGTTCTCTGAGAGCCAGTTGCGAGTA 2520
QY 808 AlaGlnProGlnAlaThrGlnProThrThrLeuAlaSerSerValAspSerAlaHisSer 827
Db 2521 GCACAGCCCAAGGTACCCAGCCGACCATTTGGCTTCTCTGTAGACAGTGCACATTTCA 2580
QY 828 AspValAlaSerGlyMetSerAspGlyAsnGluAsnValProSerSerSerGlyArgHis 847
Db 2581 GATGTTCTTCAAGTATGATGATGAGCAAGAGACCTCCATCTTCAAGTGGAGAGCAT 2640
QY 848 GluGlyArgThrThrIleArgHisTyrArgIleSerValArgSerArgSerArgHisGlu 867

```

Db	2641	GAAGGAACAATCAAAACGGCATTTACCGAAATATGTGAAGAGCGCTCTGCACATGAA	2700
Qy	868	LysThrSerArgProLysLeuArgLLeuLysnValSerAsnLysGlyAspArgValVal	887
Db	2701	AAACTTCACCCCAAAATTAAGATTTTGAATGTTTCAAAATTAAGAGACCGAGTATGTA	2760
Qy	888	GluCysGlnLeuGluThrHisAsnArgLysMetValThrPheLysPheAspLeuAspGly	907
Db	2761	GAATGTCAATTGAGACTCATATATAGGAAATAGTTGATCACTTCAAAATTTGACCTAGATGCT	2820
Qy	908	AspAsnProGlnGluLLeuAlaThrLLeuMetValAsnAsnAspPheLLeuAlaLLeuGlu	927
Db	2821	GACAAACCCGAGAGATAGCAACATTTATGTGAACATGACTTATTTCTAGCATATAG	2880
Qy	928	ArgGluSerPheValAspGlnValArgGluLLeuGluLysAlaAspGluMetLeuSer	947
Db	2881	AGAGAGTGGTTGTGTGATCAAGTCCGAGAAATTAATGAAAAAGCTGATCAAAATGCTCAGT	2940
Qy	948	GluAspValSerValGluProGlnGluLysAspGlnGluLysLeuGluLysLeuGluLysAsp	967
Db	2941	GAGGATGTCAGTGTGAACCAAGGGGTATCAAGGATTTGAGAGCTTCAAGAGAAAGAT	3000
Qy	968	AspLysGlyPheSerGlySerGlnLysLeuGluGluLysGlnProLLeuProLa	987
Db	3001	GACTATGCTTTTCAGGTTCTCAGAAATTTGGAAGAGAGTTCAAAACCAATTCCTGCG	3060
Qy	988	SerSerMetProGlnGlnLLeuGlyLLeuProThrSerSerLeuThrGlnValValHisSer	1007
Db	3061	TCTTCCATGCGCACACAAATAGCGATTCCTTACAGATTCCTTAACTCAAGTTGTTATCT	3120
Qy	1008	AlaGlyArgArgPheLLeuValSerProValProGluSerArgLysGluSerLysVal	1027
Db	3121	GCGGAGAGCGGCTTATATGTGAATCTGTCGAGAAAGCCATTAACGAAATCAAAAGTT	3180
Qy	1028	PheProSerGluLLeuThrAspThrValAlaAlaSerThrAlaGlnSerProGlyLysAsn	1047
Db	3181	TTCCCGAGTAATCAATACAGATACAGTTGCTGCTTACAGCTCAAGCCCTGGAATGAAAC	3240
Qy	1048	LeuSerHisSerLysSerSerLeuSerLeuGlnGlnAlaPheSerGluLeuArgAla	1067
Db	3241	TTGTCTCACTCTGCATCATCTCCCTTAGTCTCAACAGCGCTTTTCGAACCTTAGAGCTGCC	3300
Qy	1068	GlnMetThrGlnGluProAsnThrAlaProProAsnPheSerHisThrGlyProThrPhe	1087
Db	3301	CAATATGACGAAGAGCCCAATACAGCACTCCAACTTAGTCATACAGAGCCCAATTT	3360
Qy	1088	ProValValProProPheLeuSerSerLLeuAlaGlyValProThrThrAlaAlaThr	1107
Db	3361	CCAGTAGTACCTCTTTCTTAAGTAGCATGCTGAGAGTCCCAACCAAGCAGCAGCCACA	3420
Qy	1108	AlaProValProAlaThrSerSerProProAsnAspLLeuSerThrSerValLLeuGlnSer	1127
Db	3421	GCACAGAGTCCCTGCAACAGCAGCCCTCTATATGACATTTCCACATACGTAATTAAGTCT	3480
Qy	1128	GluValThrValProThrGlnGluGlyLLeuAlaGlyValAlaThrSerThrGlyValVal	1147
Db	3481	GAGGTTACAGTCCCACTGAAAGAGGGGATTTGCGAGTTGCCAACCAAGCAGAGTGGTA	3540
Qy	1148	ThrSerGlyGlyLeuProLLeuProProValSerGluSerProValLeuSerSerValVal	1167
Db	3541	ACTTCAGGAGTCTCCCTCCCATACACACTGTGCTGATACCAAGTACTTCCAGCGTAGTT	3600
Qy	1168	SerSerLLeuThrLLeuProAlaValValSerLLeuSerThrThrSerProSerLeuGlnVal	1187
Db	3601	TCAATATACACAATACCTGCACTGTCTCAATATCTACTACATCCCGCTCACTTCAAGTC	3660
Qy	1188	ProThrSerThrSerGluLLeuValValSerSerThrAlaLeuLysProSerValThrVal	1207
Db	3661	CCCAATTCACATCTGAGATGCTGTTTCTTAGTACAGCACTTAATCCTTCAAGTAAACAGTT	3720
Qy	1208	SerAlaThrSerAlaSerAlaGlySerThrAlaThrProGlyProLysProProAla	1227
Db	3721	TCAGCAATTCAGCCTCTGACAGGGGAGTACTGCTACCCAGAGTCTTAAGCTCCAGCT	3780
Qy	1228	ValValSerGlnGlnAlaAlaGlySerThrThrValGlyAlaThrLeuThrSerValSer	1247
Db	3781	GTAGATCTCAGCGAGGAGCGAGGAGCATCTGAGGAGGACCAATTACATCAATTTCT	3840
Qy	1248	ThrThrThrPheProSerThrAlaSerGlnLeuSerLLeuGlnLeuSerSerThrThr	1267
Db	3841	ACCACCACTTATCTCCAGACAGCTTCCAGAGCTGCTTCATTCAGCTTACAGCAGTACT	3900
Qy	1268	SerThrProThrLeuAlaGluThrValValSerAlaHisSerLeuAspLysThrSer	1287
Db	3901	TCTACTCTTACTTAGCTGAACCGTGTAGTACCGACACTCACTAGATAGACATCT	3960
Qy	1288	HisSerSerThrThrGlnLLeuAlaPheSerLeuSerAlaProSerSerSerSerPro	1307
Db	3961	CATAGCACTACAACTGATGAGCTTTCCTCTCTCTGACCATCTTCTCTCTCTCTCT	4020
Qy	1308	GlyAlaGlyValSerSerThrLLeuSerGlnProGlyLLeuHisProLeuValLLeuPro	1327
Db	4021	GGAGCAGAGTGTCTAGTATATTTCTCAGCTGTGGGCTGTCATCTTTGTGATTTCCA	4080
Qy	1328	SerValLLeuAlaSerThrProLLeuProGlnAlaAlaGlyProThrSerThrProLeu	1347
Db	4081	TCAGTAGATGCTTACTTCTTATCTTCCCAAGCAGAGACCTTACTTACACCTTTA	4140
Qy	1348	LeuProGlnValProSerLLeuProProLeuValGlnProValAlaAsnValProAlaVal	1367
Db	4141	TTACCCCAAGTACCTAGTATCCACCTTGATACGCTGTGTGCAAAATGCTGCTGTA	4200
Qy	1368	GlnGlnThrLeuLLeuHisSerGlnProGlnProAlaLeuLeuProAsnGlnProHisThr	1387
Db	4201	CAGCAGACATATTCATATAGTCAGCTCACACAGCTTTGCTTCCCAACAGCCCATCT	4260
Qy	1388	HisCysProGlnValAspSerAspThrGlnProLysAlaProGlyLLeuAspAspLLeuLys	1407
Db	4261	CATTGCTCTGAAGTAGATTCGATACACAAAGCTCTGGAATGATACATTAAG	4320
Qy	1408	ThrLeuGlnGluLysLeuArgSerLeuPheSerGluHisSerSerSerGlyValAlaGlnHis	1427
Db	4321	ACTTAGAAGAAAGCTGGCTCTGTCAGTGAACAGCTTCAATCTGAGACTAGCAT	4380
Qy	1428	AlaSerValSerLeuGluThrSerLeuValLLeuSerThrValThrProGlyLLeuPro	1447
Db	4381	GCCTCTGCTCAGTGGAGACCTCACTAGTCATAGAGACATGTCACACCGAGCATCCA	4440
Qy	1448	ThrThrAlaValAlaProSerLysLeuLeuThrSerThrThrSerThrCysLeuProPro	1467
Db	4441	ACTACTGCTGTGGCACCAAGCAAACTCCTGACTTCTACACAAAGTACTGTCTTACACCA	4500
Qy	1468	ThrAsnLeuProLeuGluThrValAlaLeuProValThrProValValThrProGlyGln	1487
Db	4501	ACCAATTTACCACTAGGAACAGTGTCTTGGCAGTTACACAGATGCTACACTGGGCA	4560
Qy	1488	ValSerThrProValSerThrThrThrSerGlyValLysProGlyThrAlaProSerLys	1507
Db	4561	GTTTCTACCCAGTACAGCACTACATCAGAGAGAAACCTGGAACCTGCTCCCTCAAG	4620
Qy	1508	ProProLeuThrLysAlaProValLeuProValGlyThrGluLeuProAlaGlyThrLeu	1527
Db	4621	CCACTCTTACTTAAGCTCCGCTGCTGCCAGGGGTACTGAACTTCCAGCAGTACTCTA	4680
Qy	1528	ProSerGlnGlnLeuProProPheProGlyProSerLeuThrGlnSerGlnGlnProLeu	1547
Db	4681	CCAGCGAGAGCTCCACCTTTTCCAGGACCTTCTTAACCACTCCACAGCACTCTTA	4740
Qy	1548	GluAspLeuAspAlaGlnLeuArgArgThrLeuSerProGluLLeuThrValThrSer	1567
Db	4741	GAGGATCTTGAATGCTCAATGGAAGAAACATTATCCAGAGATATATACAGTACTCT	4800
Qy	1568	AlaValGlyProValSerMetAlaAlaProThrAlaLLeuGluAlaGlyThrGlnPro	1587
Db	4801	GCGGTTGCTCTGTCTCATGCGGCTCCCAACAGCAATCACAGAAAGCAGGACAGGCT	4860

QY	1588	GlnlysglyValSerGlnValIlysglyGlyProValIleuAlaThrSerSerGlyValGly	1607
Db	4861	CAGAGGGGTGTTCTTCAAGTCAGAAAGAGGCCCTGCTCTAGCAACTAGTTCAGAGCTGGT	4920
QY	1608	ValPheIysMetGlyAArgPheGlnValSerValAlaIAspGlyValGlnIlysglyGly	1627
Db	4921	GTTTTTAAGATGGAGCAGATTCAAGGTTCTGTTGAGAGAGCGGTGCCAGAAAGGGGT	4980
QY	1628	LysAsnIysSerGluAspAlaIAspSerValHisPheGlySerSerThSerGlySerSer	1647
Db	4981	AAAAATAAGTCAGAAAGATGCAAAAGCTGTTCATTTTGAATTCAGACCTCAGAGCTCTCA	5040
QY	1648	ValIeuSerSerSerSerProGlySerThrIleuValIysProGluProAsnGlyIleThr	1667
Db	5041	GTGCTATCAATGATAGTACGCAGAGAGTACTTGTTGAAACCAAGACCGGATGGCATAAC	5100
QY	1668	IleProGlyIleSerSerAspValProGlySerAlaHisIlyThrThrAlaSerGluAla	1687
Db	5101	ATCCCTGGATATCTCTTCAGATGTGCAAGAGTGCACAAAACTACAGCTCAGAGCA	5166
QY	1688	LysSerAspThrGlyGlnProThrIlyValGlyAArgPheGlnValThrThrThAlaAsn	1707
Db	5161	AAGTCAGACATGGGCACTTACCAAGTTGGACTTTTCAGGTGACAACTACAGCAAC	5220
QY	1708	LysValGlyAArgPheSerValSerIysThrGluAspIlyleThrAspThrIlyIysGlu	1727
Db	5221	AAAGGGGTCTTCTCTCTGTATCAAAAACCTAGAGCAAGATCACTGACACAAAGAAAGAA	5280
QY	1728	GlyProValIAserProProPheSerPheIleuGlnAlaValIleProAlaValIle	1747
Db	5281	GGACCAAGGGATCTCCCTCTTTATGGAATTTGAAACAAGCTGTTCTTCGCTGTGTA	5340
QY	1748	ProIlyIysGlyIysProGluIleuSerGluProSerHisIleuAsnGlyProSerSerAsp	1767
Db	5341	CCAAAGAAAGAGAAAGCTTGAACGTGCAAGCTTCACTTAAATGGGCGGTCTTGAC	5400
QY	1768	ProGluAlaAlaPheIeuSerAArgAspValAspAspGlySerGlySerProHisSerPro	1787
Db	5401	CCGGAGGCGCTTTTATTAAGTAGGAGATGTGATGATGTTCCGGTAGTCCACACTCGCC	5466
QY	1788	HisGlnIleuSerSerIysSerIleuProSerGlnAsnIleuSerGlnSerIeuSerAsnSer	1807
Db	5461	CATAGCTGAGACTCAAAAGCGCTTCTTAGCCAGATCTCAAGCTCAAAAGCTTAGTAATCCA	5520
QY	1808	PheAsnSerSerIyYrMetSerSerAspAsnGlySerAspIleGlyAspGluAspIleuIys	1827
Db	5521	TTTAACTCTCTTACATGATGAGGACATAGATGATGATGATGAGATGAAGACTTAAAG	5580
QY	1828	IeuGlnIleuAArgIleuAArgIleuAArgIlyHisIleuIysGlnIleGlnAspIleuGlnSerA	1847
Db	5581	TTAGAGGCTGGCAGCGACTACGATGAATGAATCTCAAAAGAGATTCAAGACTCCAGAGTGC	5640
QY	1848	GlnIlyHisGlyIleGlySerIeuYrThrIysIleuGlyIysValProProAlaValIle	1867
Db	5641	CAGAGCATGAATATGATCTTTGTATCCAAACTGGGCAAGGTGCCCTCGCTGTTATT	5700
QY	1868	IleProProAlaAlaProIleuSerGlyAArgAArgAArgProIlyIysSerIysGlySer	1887
Db	5701	ATTCCTCCAGCTGCTCCCTTCACGGGAGAAAGACCAACCTCAAAAGCAAGGCAAGC	5766
QY	1888	LysSerSerAArgSerSerSerIleuGlyAsnIlySerProGlnIleuSerGlyAsnIleuSer	1907
Db	5761	AAATCTAGTGAAGACAGTCTCTTGGGGAATAAAGCCCGACGTTTACGTAACCTGCT	5820
QY	1908	GlyIleuSerAlaAlaIAserValIleuHisProGlnGlnThrIleuHisProProGlyAsnIle	1927
Db	5821	GGTCAGAGTGAAGCTTCAAGCTTTCACCCCAAGACAGACCTCCACCTCTCGGAACATC	5880
QY	1928	ProGlySerGlyIleAsnGlnIleuIleuGlnProIleuIysProAspProSerSerSerAspAsn	1947
Db	5881	CCAAGATCTGGGCAAGATCACTGTTACGCGCCCTTAAAGCCATCTCCCTCAGAGACAC	5940

QY	1948	Leu1yrsSer1aPheThrsSerArgG1yAla1leSerValProSerLeuSerAlaProGly	1967
Db	5941	CTCAATTCAGCCTTCACCAAGTAGGAGGCCATTTCAGTACCAAGCCTTCTGCTCCAGCT	60000
QY	1968	GlnGlyThrsSerSerThrsAnthValG1yAlaThrValAsnSerGlnAlaGlnAla	1987
Db	6001	CAAGGAACACAGCACACAAACACTGTTGGGGCAACGTGAACAGCAAGCCGCCAACCT	60606
QY	1988	GlnProProAlaMetThrSerSerArgIysGlyThrPheThrAspAspLeuHisIysLeu	2007
Db	6061	CAGCCTCCTGGCAACGACGTCACAGAGAAAGGGCACATTCACAGATGACTTCGACAAACTTG	61200
QY	2008	ValAspAsnTPAlaArgAspAlaMetAsnLeuSerGlyArgG1ySerIysGlyHis	2027
Db	6121	GTAACCAATTTGGGCCGAGATGCGCATGATCTTCAGGACGGAAGAGAAAGGAGGAC	61800
QY	2028	MetAsnTPArgGlyProGlyIysMetAlaArgIysPheSerAlaProGlyGlnLeuCysIle	2047
Db	6181	ATGAATATATAGAGGGCCCTGGAAATGGCAAGGAAGTTCTCTGCACTGGGCAACTGTGACTC	62400
QY	2048	SerMetThrsSerAsnLeuGlyIysSerAlaProIleSerAlaAlaSerAlaThrSerIleu	2067
Db	6241	TCCATGACCTCGAAGCTGGGGTGGCTCTGCCCCCACTCTCTGACGATCAGCTAACCCTCTA	63000
QY	2068	GlyHisPheThrIysSerMetCysProProGlnGlnIlyrGlyPheProAlaThrProPhe	2087
Db	6301	GGTCACTTCACCAAGCTCTATGTGCCCCCACAGAGATAGGCTTCCAGTACCCCACTT	63600
QY	2088	GlyAlaGlnTPSerSerGlyThrGlyIyProAlaProGlnProLeuGlyIyGlnPheGlnPro	2107
Db	6361	GGCGCTCATGATGGATGGGACCGGGTGGCCGACACACAGCCACTTGGCCAGTTCCAACT	64200
QY	2108	ValGlyThrAlaSerLeuGlnAsnPheAsnIleSerAsnLeuGlnIysSerIleSerAsn	2127
Db	6421	GTGGGAACCTGGCTCCTTGACGAATTTCAACATCAGCAATTTGACGAATATCATCAGCAAC	64800
QY	2128	ProProGlySerAsnLeuArgThrThr	2136
Db	6481	CCCCAGGCTCCAACTGGCGGACCACT	6507
RESULT 11			
LOCUS	AB002342	6812 bp	mRNA linear PRI 10-MAY-2002
DEFINITION	Human mRNA for KIAA0344 gene, partial cds.		
ACCESSION	AB002342		
VERSION	AB002342.2	GI:20521006	
KEYWORDS	KIAA0344.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	1		
AUTHORS	Nagase,T., Ishikawa,K., Nakajima,D., Ohira,M., Seki,N., Miyajima,N., Tanaka,A., Kotani,H., Nomura,N. and Ohara,O.		
TITLE	Prediction of the coding sequences of unidentified human genes. VII. The complete sequences of 100 new cDNA clones from brain which can code for large proteins in vitro		
JOURNAL	JOURNAL OF MOLECULAR BIOLOGY		
REFERENCE	2 (bases 1 to 6812)		
AUTHORS	Ohara,O., Nagase,T., Kikuno,R. and Nomura,N.		
TITLE	Direct Submission		
JOURNAL	Submitted (28-MAR-1997) Osamu Ohara, Kazusa DNA Research Institute; 1532-3, Yana, Kisarazu, Chiba 292-0812, Japan (E-mail:cdnainfo@kazusa.or.jp, Tel:+81-438-552-3913)		
COMMENT	On May 9, 2002 this sequence version replaced gi:2280478. Sequence updated (22-Jul-1997).		
FEATURES	Location/Qualifiers		
SOURCE	1..6812		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		

```

/clone="hg01568"
/sex="male"
/cisue_type="brain"
/clone_lib="pbluescriptII SK plus"
/notes="this sequence was replaced that of hg01486 cDNA as
a representative cDNA sequence for KIAA0344."
1..6812
/gene="KIAA0344"
<1..6203
/gene="KIAA0344"
/codon_start=3
/protein_id="BAA20802.2"
/db_xref="gi:20521007"
/translation="AATTTTTHRRFRSVCDSNATLRLGLPLSLPSPISPAV
OSAPRPRRTVATATTSOVAOOPRAAREGONAPRATVTSSTSSKSPKSPAS
LVGSEKPPRPSAGSGGSAKPEERQODDIELTAKVGNDRFLKFDIEIG
KGSFTVKIGLDTETIVAMCELDRLTKRSERQFEEMKLGQHPNIVRYS
MESTVKKKIVLVTMLTSGTLKTKYKFKMKIKVISMCRQILKGLQELRTP
IHRDKCDNIFITGPTGSVKIGLGLATLKRASPAKSVIGPHEMAREBEKES
VDVAFKGMCEMATSEYSPSCQNAQIYRVTSGVADKXVAIPEVEIEGCI
RONKDERYSKIDLNHAFPOERTGVRLAEEDDEKAIKILMRIDENIKLKGVKD
NEAIFSFDLERDVEDVAQEMVSEGYCEBQHKMAKIDRSLIRKAKRQLVR
BEQKKQESLKOQVQSSASQTGIQLPSASTGTPASTTSASVSTOVEPEEA
DOHQLOQVQPSISVLSDEIVDSQGSQSVFESRVSQQTYSQHQEQASTGVP
HISTVQAQSPHGVPPSVAAQSQSQSSSLTGSSQPIQHPQOQCIQCTAP
OOTVOYLSQTSSEATTAQOPPOVLPVYAGKQSTQGVQVAPAPVAVAO
POATOPTLASSVDSASHDVAQMGDMENVPSSGRREGRTTKHYKXSVRSRHE
KTSRKLAILVSNKGRVVEQCLETHNRKAVTFKFDLDGNPEIATIMNNDTLA
TERESFVQVEEIIKADEMLSESVPEEQDQLESIQGKODVFSQQLKEFEKQ
PIPASMPQOIGIPITSLTOVHSAARFIVSPESRLREKVPSEITDTVASTA
OSPKNLSHSAISLSLQAFSELRAQTEGNAPPMHSTGPTFPVPVPELISIA
VPTTAATAPVATSSPNIDISTVISEVTPTEGAGVATSGVATSGVLPV
SESPLSSVSSITIPAVYISTSPISOVPTSEIYSSVATPSVTVATSTASAG
GSTAPGKPRPAVVSQAAGSTTVATITSTSTTSFPTASQLSIQSSSTSTTTLA
ETVVASHSLSLDTSHSSTGLAFSLASVSSSSPGAVSYISQPGALPLVTSVIA
STPLIPQAAGPTSTPLPQVPSIPLVQVNAVPOQLHISQPGALPLVPTBVA
PEVSDTQPKAAGIDIDIKLEKLSLSHSSGCAHSAVLSLSTVETGIP
TTAVASKLSTSTSTCLPPTNLPLGTVALPVTPVTPGOVSTPVSTTSSGKPTAP
SKPPLTKAPLVPGTELPAQTLPSROLPPFPGPSLTOSQPLEDLDRLTISPEM
TVTSRVGVSMAAPPAITEAGPQKVSQVKEGPLATSSGAVGKMRQVSVAAD
AOKGKNSSEDAKSHFESSTESSVSSSESTLVPEPGLITPISSDVPSAH
KTJASEASDQPTKGRFOVTTTANVGRSVSKTDITDKKEBPSPMDL
BOAVLPAVTPKKEKELSEPSHLNDSDEPAAPLSRDVDSGPHOSLSKSLP
SONLSQSLSNSNSNSNDSEIDEDLDKLELRDKLKEIJDOSQKHIES
LYTKIGKVPAYIIPPAISGRPRPKSGKSGKSRYSAPFSDAKISPSLSGQSA
SVLHCKTLHPRPNIPESGNDLLOPLKPSRSDKYSAPFSDAKISPSLSGQSA
SSTNVGATVNSQAQAPPAWTSRKSTFTDHLKLVNARDMNLSGRSGGHM
NYEGMARKSAPQOLCI SMTSNLGSAPISAAATSI GHTKSMCPPOYGFPAFP
FGAWSGTGPAFQGFQFQPVGTASLQNFNISNLQKSI SNPPGSLRRT"

```

ORIGIN

Alignment Scores:

```

Pred. No.: 1,02e-248 Length: 6812
Score: 10435.00 Match: 2065
Percent Similarity: 99.85% Conservative: 0
Best Local Similarity: 99.85% Mismatches: 1
Query Match: 96.51% Indels: 2
DB: 9 Gaps: 2

```

US-10-010-720-14 (1-2136) x AB002342 (1-6812)

```

QY 69 AAlaIaThrThrThrGluHisArgPheArgSerValIleCysAspSer 88
Db 3 GCGGGAGACCACTACCACTAGACACCGCTTCCGCCGAGGCTCATCTGCACATCC 62
QY 89 AenAlaThrAlaLeuGluLeuProGlyLeuProLeuSerLeuProGlnProSerIlePro 108
Db 63 AATGCACATGCAGCTGAGCTTCCCGGCTTCTTCTTCCCTGCCCCAAGCCAGCATCC 122
QY 109 AAlaIaValProGlnSerAlaProProGluProHisArgGluGluThrValThrAlaThr 128
Db 123 GCGGCTGTCCCGCAGAGTCTCCACCGGAGCCCAACCGGAGAGACCGTGAACCCGAC 182

```

```

QY 129 AAlaThrSerGlnValAlaGlnGlnProProAlaAlaAlaProGlyGluGlnAlaVal 148
Db 183 GCACTTCCAGAGTACAGCAGAGCTCAGAGCTGCGCCCTTGAGGAAACAGGCGCTC 242
QY 149 AAlaGlyProAlaProSerThrValProSerSerThrSerIlyAspArgProValSerGln 168
Db 243 GCGGGCCCTGCGCCCTGCACTGTCTCCAGCAGTACAGCAAAAGACCGCCAGTGTCCAG 302
QY 169 ProSerLeuValGlySerIlyGluGluProProProAlaArgSerGlySerGlyGly 188
Db 303 CCTACCTTGTGGGAGAGCAAGAGAGCGCGCGGAGAGAGAGGAGCGCGCGCGCGC 362
QY 189 SerAlaIyGluProGlnGluGluArgSerGlnGlnAspAlaGluGluLeuGlu 208
Db 363 AGCCGCAAGAGCCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 422
QY 209 ThrIlyAlaValGlyMetSerAsnAspGlyArgPheLeuIlyGluIleGly 228
Db 423 ACCAAGCGCGGAGAGATGCTTACAGATGCGCGCTTCTCAAGTTTGACATCGAAATCG 482
QY 229 ArgGlySerPheIlyThrValTyrIlyGlyLeuAspThrGluThrValGluValAla 248
Db 483 AGAGGCTCTTTAAGACGCTTACAAAGGCTTGACACTGAAACCAAGGAGAGAGTCCGC 542
QY 249 TrpCysGluLeuGlnAspArgIlySerThrIlySerSerGluArgGlnArgPheIlyGlu 268
Db 543 TGGTGTGAACGCGAGATCGAAATTAACAAAGTCTGAGAGGCGAGAGTTTAAAGAA 602
QY 269 AAlaGluLeuIlyGlyLeuGlnHisProAsnIleValArgPheTyrAspSerTrpGlu 288
Db 603 GCTGAAAGTTTAAAGGCTTACAGATCCCATATATGTTAATTTATGATTTCCGGGAA 662
QY 289 SerThrValIlyGlyIlySerCysIleValLeuValThrGluLeuMetThrSerGlyThr 308
Db 663 TCCACAGTAAAGAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 722
QY 309 LeuIlyThrTyrLeuIlyArgPheIlyValMetIlyIlyValLeuArgSerTrpCys 328
Db 723 CTTAATAAGTATCGAATAAGGTTTAAATGATGAAAGTCTTAAGAGAGAGAGAGAG 782
QY 329 ArgGlnIleLeuIlyGlyLeuGlnPheLeuHisThrArgThrProProIleIleHisArg 348
Db 783 GGTGAGATCTCTTAAAGGCTTACAGATCTTCAATCTCAACCTCAATCAATCAATCA 842
QY 349 AspLeuIlyCysAspAsnIlePheIleThrGlyProThrGlySerValIlyIleHisArg 368
Db 843 GATCTTAATGTGACAAATCTTTATCAACCGGCTTACTGCTCAGTCAAGATGTGAGAC 902
QY 369 LeuGlyLeuAlaThrLeuIlyArgAlaSerPheAlaIlySerValIleGlyThrProGlu 388
Db 903 CTCGGCTCGGCAACCTTGAAGGAGGCTTCTTTGCGAAGAGTGTATGTATGCCCAAG 962
QY 389 PheMetAlaProGluMetTyrGluGluIlyIlyIlyIlyIlyIlyIlyIlyIlyIly 408
Db 963 TTCAATGAGGCTTGAATGATGAGAGAGAAATATGATGATCCGTGAGAGTATATCTTT 1022
QY 409 GlyMetCysMetLeuGlnMetAlaThrSerGluTyrProTyrSerGluCysGlnAsnAla 428
Db 1023 GAGATGTCATGCTTGAATGCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTT 1082
QY 429 AAlaGlnIleTyrArgArgValThrSerGlyValIlyIlyProAlaSerPheAspIlyValAla 448
Db 1083 GCGCAGATCTACCGTCCGCTGACAGTGGGGGAGAGCCAGAGTTTGAACAAATGACA 1142
QY 449 IleProGluValIlyGluIleIleGluGlyCysIleIleGlnAsnIlyAspGluArgTyr 468
Db 1143 ATTCCTGAAGTGAAGAAATATGAGAGATGACATACCAAAAGAGAGAGAGATAT 1202
QY 469 SerIleIlyAspLeuLeuAsnHisAlaPhePheGlnGluGluThrGlyValArgValGlu 488
Db 1203 TCCATCAAGACCTTTTAAACATGCTTCTTCAAGAGAGAGAGAGAGAGAGAGAGAG 1262
QY 489 LeuAlaGluGluAspAspGlyGluIlyIlyIleAlaIleIlyIlyIlyIlyIlyIly 508

```

```

Db      1263 TTAGCAGAAAGAAATGATGAGCAAAAATAAGCCATTAATTAATGCTAGCTATGAAAGAT 1322
Qy      11elyslyLeuIyeglylyserTyryAspAspAngluAlaIlegluPheSerPheAspLeu 528
Db      1323 ATTAAGAAATTAAGGAAATACAAAGATTAATGAAGTATGAGATTTCTTTGATTTTA 1382
Qy      529 GluArgAspValProGluAspValAlaGlnGlnMetValGlnSerGlyTyryValCysGlu 548
Db      1383 GAGAGAGATGTCCTCCAGAAAGATGTCACAAAGAAATGATGAGTCTGGGATATGTCGTGAA 1442
Qy      549 GlAspHislyeThMetAlaIyAlaIlelyAspArgValSerLeuIlelySarGlys 568
Db      1443 GGTGATCACAGAACCAATGGCTAAAGCTATCAAAAGACAGATTCATTAATTAAGGAAA 1502
Qy      569 ArgGluGlnArgGlnLeuValArgGluGlnGlnGlnIlylyeGlnIleuPheSer 588
Db      1503 CGAGAGCAGCGGCGAGTGTGCTACGGAGAGAGCAAGAAAAAAGCAGAGAGAGCAAGT 1562
Qy      589 LeuIyeglnGlnValGlnIleuSerSerAlaSerGlnThrGlyIlelyeGlnIleuPheSer 608
Db      1563 CTCAAACAGCAGGTAGAAACAATCCAGTGTCTCCAGACAGGAATCAGACAGCTCCCTCT 1622
Qy      609 AlAserThrGlyIleProThrAlaSerThrThrSerAlaSerValSerThrGlnValGlu 628
Db      1623 GCTACACCGGCAATACCTACTGCTCTCACTCACTCACTCACTCACTCACTCACTCACTCA 1682
Qy      629 ProGluGluProGluAlaAspGlnHisGlnGlnIleuGlnIlyGlnIleProSerIleSer 648
Db      1683 CCTGAAGAACCTGAGGAGATCAACATCAACAACATCAACATCAACATCAACATCAACATCT 1742
Qy      649 ValLeuSerAspGlyThrValAlaAspSerGlyGlnIySerSerValPheThrGlnSerArg 668
Db      1743 GTGTATCTGATGGGACCGTTCAGACAGTGTGTGAGGATCTCTGCTTCAAGAAATCTCGA 1802
Qy      669 ValSerSerGlnIleThrValSerTyryGlySerGlnHisGlnGlnAlaHisSerThrGly 688
Db      1803 GTGACACCCCAACACAGATTCATATGATGATTCCTCCCAACATGAACAGGCACTTCTCAAGC 1862
Qy      689 ThrValProGlnHisIleProSerThrValGlnAlaGlnSerGlnProHisGlyValTyry 708
Db      1863 ACAGTCCAGGGCAATACCTCTTCACTCTCAAGACAGCTCTCAAGCCCAATGGGATTAAT 1922
Qy      709 ProProSerSerValAlaGlnIyGlnSerGlnIyGlnIleProSerSerSerLeuThr 728
Db      1923 CCACCCTCAAGTGTGGCAGAGGGGAGAGCCAGGGTCAAGCATCTCAAGTAGCTTAACA 1982
Qy      729 GlyValSerSerSerGlnProIleGlnHisProGlnGlnGlnIyIleGlnIleThr 748
Db      1983 GGGGTTTCACTTCCCAACCCATACACATCTCT--CAGCAGCAGGGAAATCAGACAGCA 2039
Qy      749 AlaProProGlnGlnIleThrValGlnTyrySerLeuSerGlnIleThrSerSerSerGluAla 768
Db      2040 GCCCTCTCTCAACAGACAGCTGAGTATCACTTCAACAGACATCAACTCCAGGTAGGCC 2099
Qy      769 ThrThrAlaGlnProValSerGlnProGlnAlaProGlnValIleProGlnValSerAla 788
Db      2100 ACTACTGCACAGCAGTAGTCAACCTCAAGCTCCACAAAGTCTTGCCCTCAAGTAGTACGCT 2159
Qy      789 GlyIyGlnIleThrGlnGlnIyValSerGlnValAlaProAlaGluProValAlaValAla 808
Db      2160 GGAAGAAAGAGATCTCAGAGGAGTCTCAGAGTGTCTCTGCAAGGCCAGTTGACGTAGCA 2219
Qy      809 GlnProGlnAlaThrGlnProThrThrLeuAlaSerSerValAspSerAlaHisSerAsp 828
Db      2220 CAGGCCCAAGCTAACCCAGCCACCACTTTGGCTTCTCTGAGACAGTGCATTCAGAT 2279
Qy      829 ValAlaSerGlyMetSerAspGlyAsnGluAsnValProSerSerSerGlyArgHisGln 848
Db      2280 GTTCTTCAAGGTATAGTAGGCAATGAGAACTGCCATCTTCCAGTGAAGGAGCATGAA 2339
Qy      849 GlyArgThrThrIlySarGHisTyryArgIySerValArgSerArgSerArgHisGluIly 868

```

```

Db      2340 GGAGAACTACAAAAGCGCATTTACCGAAAATCTGTAAAGAGTCCGCTTCGACATGAAAAA 2399
Qy      869 ThrSerArgProIyLeuArgIleLeuAsnValSerAsnIylySerArgValValGlu 888
Db      2400 ACTTACGCCCAAAATTAAGAAATTTGAAATGTTTAAATTAAGGAGACCGAGATGAGAA 2459
Qy      889 CysGlnLeuGlnIleHisAsnArgIyMetValThrPheIyPheAspLeuAspGlyAsp 908
Db      2460 TGTCAATTAAGAGACTCAATATGAGAAATGGTTACATTCMAATTTGACCTAGATGTGAC 2519
Qy      909 AsnProGluGlnIleAlaThrIleMetValAsnAsnAspPheIleuAlaIleGluArg 928
Db      2520 AACCCTGAGAGATGCAACAATTAATGCTGAACATGACTTATCTAGCAATAGAGAGA 2579
Qy      929 GluSerPheValAspGlnValArgGluIleIleGluIlyAlaAspGlnMetLeuSerGlu 948
Db      2580 GAGTGTGTTGGATCAAGTGGAGAAATTAATGAAAAAGCTGAGAAATGCTCGAGTAG 2639
Qy      949 AspValSerValGluProGluGlyAspGlnIyLeuGlnIyLeuGlnIylyAspAsp 968
Db      2640 GATGTCAGTGGAGAACCAAGAGGTGATCAGGGATTTGAGAGCTTCAAGAGAAAGATGAC 2699
Qy      969 TyryGlyPheSerGlySerGlnIlyLeuGlnIyGluPheIyGlnProIleProAlaSer 988
Db      2700 TATGCTTTTCAAGGTTCTCAGAAATGGAGAGAGTTCAAAACAACCAATTCCTCGCTCT 2759
Qy      989 SerMetProGlnGlnIleGlyIleProThrSerSerLeuThrGlnValValHisSerAla 1008
Db      2760 TCCATGCCACAGCAATATGGCAATTCCTACAGTCTTCACTCAAGTGTGATTCCTGG 2819
Qy      1009 GlyArgArgPheIleValSerProValProGluSerArgLeuArgGluSerIyValPhe 1028
Db      2820 GGAAGCGGTTATATAGTAGTCTGTGCCAAGAACCGATTAACAGAAATCAAAATTTTC 2879
Qy      1029 ProSerGluIleThrAspThrValAlaAlaSerThrAlaGlnSerProGlyMetAsnLeu 1048
Db      2880 CCCAGTGAATTAACAGATACAGTGTGCTGCTCTCACTCACTCACTCACTCACTCACTG 2939
Qy      1049 SerHisSerAlaSerSerLeuSerLeuGlnIlePheSerGluIleArgArgAlaGln 1068
Db      2940 TCTCACTGTGATCATCTCTTAAGTACATCAACAGGCTTTTGTGAACCTTAAGACGTGCCAA 2999
Qy      1069 MetThrGlnIyProAsnThrAlaProProAsnPheSerHisThrGlyProThrPhePro 1088
Db      3000 ATGACAGAGAGCAACCAATACAGACACCTCCAACTTGTACTATACAGACCAACATTTCCA 3059
Qy      1089 ValValProProPheLeuSerSerIleAlaGlyValProThrThrAlaAlaIleThrAla 1108
Db      3060 GTAGTACCTCTCTTCTTAAGTAGCATTTGCTGAGATGCCAACACAGCAGCAGCACAGCA 3119
Qy      1109 ProValProAlaThrSerSerProProAsnAspIleSerThrSerValIleGlnSerGlu 1128
Db      3120 CCAATCCCTGCAACAGAGGCCCTCTTAATGATCAATTTCCACATGATGATTAATCACTGTAG 3179
Qy      1129 ValThrValProThrGlnGlnIyIleAlaGlyValAlaThrSerThrGlyValValThr 1148
Db      3180 GTTACAGTGGCCCACTGAAGAGGGGATGCTGAGATTCACACACACAGAGTGTGTAAT 3229
Qy      1149 SerGlyGlyLeuProIleProProValSerGlnSerProValLeuSerSerValValSer 1168
Db      3240 TCAGTGTGTTCCCAATACCACTGTCTGAATCAACAGATCACTTCCAGGTGTGTTCA 3299
Qy      1169 SerIleThrIleProAlaValAlaSerIleSerThrThrSerProSerLeuGlnValPro 1188
Db      3300 AGTATCCAAATACCTGCAAGTGTGCTCAATATCTACATCCCGGTCACTTCAAGTCCCC 3359
Qy      1189 ThrSerThrSerGlnIleValAlaSerSerThrAlaLeuTyryProSerValThrValSer 1208
Db      3360 ACATCCACATCTGAGATCGTTGTTTCTAGTACAGACATGATCTCTTCAGTAAACAGTTTCA 3419
Qy      1209 AlaThrSerAlaSerAlaGlyGlySerThrAlaThrProGlyProIyProProAlaVal 1228
Db      3420 GCAACTTCAAGCTCTGAGGGGAGATGCTAATGCCCAAGTCTTAAGCTTCAAGCTGTA 3479

```


QY 1229 ValSerGlnGlnAlaGlySerThrThrValGlyAlaThrLeuThrSerValSerThr 1248
DB 3480 GTATCTCGACGAGCAGCAGCAGCAGCTGTGGAGCCACATTACAGTTCTTACCC 3539
QY 1249 ThrThrSerPheProSerThrAlaSerGlnLeuSerIleGlnLeuSerSerThrSer 1268
DB 3540 ACGACTTATTCACAGCAGCAGCTTCAAGCTGTCCATTCAAGCTTACAGCAGTACTTCT 3599
QY 1269 ThrProThrLeuAlaGlnThrValValSerAlaHisSerLeuAspLysThrSerHis 1288
DB 3600 ACTCTACTTATGAGTGAACCGGTGAGTTCGCGACACTCACTAGATTAAGACATCTCAT 3659
QY 1289 SerSerThrThrGlyLeuAlaPheSerLeuSerAlaProSerSerSerSerProGly 1308
DB 3660 AGCGATCAACTGATGATGAGCTTCTCTCTCTGACACACTTCTCTCTCTCTCTCTG 3719
QY 1309 AlaGlyValSerSerThrIleSerGlnProGlyGlyLeuHisProLeuValIleProSer 1328
DB 3720 GCAGGAGTGTCTAGTTATATTTCTCAGCCTGGTGGCGCATCTTGTGTCATTCATCA 3779
QY 1329 ValIleAlaSerThrProIleLeuProGlnAlaAlaGlyProThrSerThrProLeu 1348
DB 3780 GTGATAGCTTCTACTCTTATTTCTCCCAAGCAGCAGACCTACTTCTTATTA 3839
QY 1349 ProGlnValProSerIleProProLeuValGlnProValAlaAsnValProAlaValGln 1368
DB 3840 CCCCAAGTACTGATGATTCACCCCTGTGACAGCTGTGCAATGTCTGTCTGACG 3899
QY 1369 GlnThrLeuIleHisSerGlnProGlnProAlaLeuLeuProAsnGlnProHis 1388
DB 3900 CAGACACTATTCATAGTACAGCTTCAACAGCTTGTCTTCCCAACAGCCCATCATCT 3959
QY 1389 CysProGlnValAspSerAspThrGlnProLysAlaProGlyIleAspAspIleLysThr 1408
DB 3960 TGTCTGGAAGTATTCGATTCACACCAACCAAGCTCTGGAATGATGACATAAAGCT 4019
QY 1409 LeuGlnGlyLysLeuArgSerLeuPheSerGlnHisSerSerSerGlyAlaGlnHisAla 1428
DB 4020 CTGAGAGAAAGCTCGGCTCTCTGTTCACTGACAGCTCATCTGAGACTGACAGATGCC 4079
QY 1429 SerValSerLeuGlyThrSerLeuValIleGlnSerThrValThrProGlyIleProThr 1448
DB 4080 TCTGTCTCAGTGAACCTCATCTAGTACATAGAGACATGTCTACACAGGACTTCCACT 4139
QY 1449 ThrAlaValAlaProSerLysLeuLeuThrSerThrThrSerThrCysLeuProProThr 1468
DB 4140 ACTGCTGTGACACCAAGCAAACTCTGACTTACCAAGTATGCTTACCAACCAAC 4199
QY 1469 AsnLeuProLeuGlyThrValAlaLeuProValThrProValValThrProGlyGlnVal 1488
DB 4200 AATTATCACTAGAACAGCTTCTTCCAGTTTACCAAGTGTACACCTTGGGCAAGTT 4259
QY 1489 SerThrProValSerThrThrThrSerGlyValLysProGlyThrAlaProSerLysPro 1508
DB 4260 TCTACCCAGTACAGCACTCATCATCAGAGATGAACCTCGAAGTCTCTCCCAAGCA 4319
QY 1509 ProLeuThrLysAlaProValLeuProValGlyThrGlnLeuProAlaGlyThrLeuPro 1528
DB 4320 CCTTCACTAAGGCTCCGGTGTGCGCAGTGGTACTGAACTTCCACAGGATCTTACCC 4379
QY 1529 SerGlnGlnLeuProProPheProGlyProSerLeuThrGlnSerGlnGlnProGly 1548
DB 4380 ACGAGACAGCTCCACCTTCTTCCAGCACTTCTTACCACTCCAGCACTCTTAGAG 4439
QY 1549 AspLeuAspAlaGlnLeuArgThrThrLeuSerProGlu**IleThrValThrSerAla 1568
DB 4440 GATCTTGATGCTCAATTGAGGAACACTTAGTCCAGAGATGATCACTGACTTCTGGG 4499
QY 1569 ValGlyProValSerMetAlaAlaProThrAlaIleThrGlnAlaGlyThrGlnProGln 1588
DB 4500 GTTGTCTCTGTGTCAATGGCGGCTCCACAGCAATCAAGAGAGCAAGCAAGCTTAG 4559

QY 1589 LysGlyValSerGlnValLysGlyProValLeuAlaThrSerSerGlyAlaGlyVal 1608
DB 4560 AAG--GTTTCTCAAGTCAAGAAAGAGCCCTCTCTAGTACTAGTTCAGAGCTGTCTT 4616
QY 1609 PheLysMetGlyArgPheGlnValSerValAlaAlaAspGlyAlaGlnLysGlyLys 1628
DB 4617 TTTAAGATGGGACGATTTCAAGTTCTGTGTGACAGAGCGGTGCCAAGAAAGAGGTAAA 4676
QY 1629 AsnLysSerGlyAspAlaLysSerValHisPheGlnSerSerThrSerGlnSerVal 1648
DB 4677 AATAAGTCAGAAAGTGAAGTCTGTTCATTTTGAATCAGACACTCAGAGTCTCAGTG 4736
QY 1649 LeuSerSerSerSerProGlnSerThrLeuValLysProGlnProAsnGlyIleThrIle 1668
DB 4737 CTATCAAGTACTAGTCCAGAGATACCTTGTGTGAACCAAGGCCAATGGCATTAACATTC 4796
QY 1669 ProGlyIleSerSerAspValProGlnSerAlaHisLysThrThrAlaSerGlnAlaLys 1688
DB 4797 CCTGTATCTCTTCAGATGATGTCACAGAGTGCCTCAAAACTACTGCTCAGAGGCAAG 4856
QY 1689 SerAspThrGlyGlnProThrLysValGlyArgPheGlnValThrThrAlaAsnLys 1708
DB 4857 TCAAGACTGGGAGCCTTACCAAGGTTGACGTTTCAAGTGACACTACAGCAACAA 4916
QY 1709 ValGlyArgPheSerValSerLysThrGlnAspLysIleThrAspThrLysLysGly 1728
DB 4917 GTGGGTGTTTCTCTGATCAAAAAGTGAAGCAAAATCAGTCAACAAAGAAAGAA 4976
QY 1729 ProValAlaSerProProPheMetAspLeuGlnAlaValLeuProAlaValIlePro 1748
DB 4977 CCAGTGCATCTCTCTCTTATGATTTGAAACCAAGCTGTCTTCTGCTGTGATACA 5036
QY 1749 LysLysGlyLysProGlnLeuSerGlnProSerHisLeuAsnGlyProSerSerAspPro 1768
DB 5037 AAGAAAGAAAGCTGAACCTGACAGCTTCAACCTTCAATAGGCGCTTCTTCAACCG 5096
QY 1769 GlnAlaAlaPheLeuSerArgAspValAspAspGlySerGlySerProHis 1788
DB 5097 GAGGCGCTTTTATAGTAAAGGATGTGATGATGTTCCGGTATGCCACATCGCCCAT 5156
QY 1789 GlnLeuSerSerLysSerLeuProSerGlnAsnLeuSerGlnSerLeuSerAsnSerHe 1808
DB 5157 CAGCTGAAGCTCAAAAGAGCTTCTTACAGCAATCAAGTCAAAAGCTTATGATTCATT 5216
QY 1809 AsnSerSerLysMetSerSerAspAsnGlnSerAspIleGlnAspGlnAsnLeuSerLeu 1828
DB 5217 AACTCTCTTACATGATGATGACACATGATGATGATGATGATGATGATGATGAT 5276
QY 1829 GlnLeuArgArgLeuArgAspLysHisLeuLysGlnIleGlnAspLeuGlnSerArgln 1848
DB 5277 GAGCTGCAGCAGCATCAGATTAACATCTCAAAAGATTCAGGACTCGCAGAGTCCGAC 5336
QY 1849 LysHisGlnIleGlnSerLeuThrThrLysLeuGlyLysValProProAlaValIleIle 1868
DB 5337 AAGCATGAATTTGATCTTTGATACCAAACTGGGCAAGGTGCCCTGCTGTATTAAT 5396
QY 1869 ProProAlaAlaProLeuSerGlyArgArgArgProThrLysSerLysGlySerLys 1888
DB 5397 CCCCAGCTGCTCCCTTTCAGGAGAAAGACACACACCACTTAAGCAAAAGCAGCAAA 5456
QY 1889 SerSerArgSerSerSerLeuGlyAsnLysSerProGlnLeuSerGlyAsnLeuSerGly 1908
DB 5457 TCTATCGAAGAGTTCCTTGGGGAATTAAGCCCCAGCTTTCAGTACCTGCTGTGT 5516
QY 1909 GlnSerAlaAlaSerValLeuHisProGlnGlnThrLeuHisProProGlyLysAsnIlePro 1928
DB 5517 CAGAGTGCAGCTTCACTTTCAGCCCCCAGCAGACCTTCCACCTCTGGCAACATCCA 5576
QY 1929 GlnSerGlyGlnAsnGlyLeuLeuGlnProLeuLysProSerProSerSerAspAsnLeu 1948
DB 5577 GAGTCCGGGCAAAATCGCTGTTACAGCCCTTAAGCCATCTCTCCCTCAGTACAACTTC 5636
QY 1949 TyrSerAlaPheThrSerAspGlyAlaIleSerValProSerLeuSerAlaProGlyGln 1968


```

Db      5637 TATTGAGCTTCAACGATGATGCTGCTTTCAGTACCAAGCTTTGCTGCTCAGGTCAA 5696
Qy      1969 GJYThrSerSerThraSerThrValGlyAlaThrValAsnSerGlnAlaIleGlnAlaGln 1988
Db      5697 GGAACACGACGACAAACACTGTTGGGCGACAGTGAACAGCCAGCCGCCCAAGCTCAG 5756
Qy      1989 ProProAlaMetThrSerSerAlaGlyGlyThrPheThrAspAspLeuHisIleVal 2008
Db      5757 CCTCTGCGCATGACGCTCCAGCAGAGGCGACATTCACAGATGACTTGCAACAAGTTGGTA 5816
Qy      2009 AspAsnThrAlaArgAspAlaMetAsnLeuSerGlyArgArgGlySerIleGlyHisMet 2028
Db      5817 GACAAATTGGGCCCGAGATGCCATGATCTCTCAGGCGAGAGAGAAAGGACAAAGCGACAG 5876
Qy      2029 AsnIleGlyGlyProGlyMetAlaArgIleGlyPheSerAlaProGlyGlnLeuCysIleSer 2048
Db      5877 AATTATGAGGGCCCTGGATATGGCAAGAAATTCTCTGCACTGGGCACTGTGACTCTCC 5936
Qy      2049 MetThrSerAsnLeuGlyGlySerAlaProIleSerAlaIleSerAlaThrSerLeuGly 2068
Db      5937 ATGACCTCGAACCTGGGTGGTCTCTGCCCCATCTCTGCAGCATCAGTACCTCTCTAGGT 5996
Qy      2069 HisPheThrIleSerMetCysProProGlnGlnIleIleIleIleIleIleIleIleIleIle 2088
Db      5997 CACTTCACCAAGTCTATGTGCCCCCAGCAGAGTATGGCTTCCAGTACCCCATTTGGC 6056
Qy      2089 AlaGlnIlePheSerGlyThrGlyGlyProAlaProGlnProLeuGlyGlyIlePheGlnProAl 2108
Db      6057 GCTCAATGAGAGTGGGAGCGGGTGGCCAGACACACAGCCACTTGGCAGTCTCCAGCTGTG 6116
Qy      2109 GlyThrAlaSerLeuGlnAsnPheAsnIleSerAsnLeuGlnIleIleIleIleIleIleIle 2128
Db      6117 GGAACCTGCTCTCTTCAGATTTTCAACATCAGCAATTGGCAAAATTCATCAGCAACCC 6176
Qy      2129 ProGlySerAsnLeuArgThrThr 2136
Db      6177 CAGGCTCCACCTCGGACCACT 6200

```

RESULT 12

```

AR302604 6969 bp DNA linear PAT 12-JUN-2003
LOCUS AR302604
DEFINITION Sequence 33 from patent US 6541252.
ACCESSION AR302604
VERSION AR302604.1 GI:31690902
KEYWORDS
SOURCE
ORGANISM
REFERENCE
  1 (bases 1 to 6969)
  Walke, D.W., Hilbun, E., Donoho, G. and Turner, C.A. Jr.
  Human kinases and polynucleotides encoding the same
  Patent: US 6541252-A 33 01-Apr-2003;
  Location/Qualifiers
  1: 6969
  /organism="unknown"
  /mol_type="genomic DNA"

```

ORIGIN

```

Alignment Scores:
Pred. No.: 3.15e-247 Length: 6969
Score: 10375.00 Matches: 2076
Percent Similarity: 89.41% Conservative: 0
Best Local Similarity: 89.41% Mismatches: 0
Query Match: 95.96% Indels: 246
DB: 6 Gaps: 1

```

US-10-010-720-14 (1-2136) x AR302604 (1-6969)

```

Qy      61 MetAspIysAspSerArgGlyAlaIleAlaThrThrThrThrGluHisArgPhePhe 80
Db      1 ATGGACAAAGACAGCGCGTGGGGCGCGCCGACCACTACCACTGAGCAACGCTTCTTC 60

```

```

Qy      81 ArgArgSerValIleCysAspSerAsnAlaThrAlaLeuGluLeuProGlyLeuProLeu 100
Db      61 CGCCGAGAGTGCATCTGGAGATCCATGCACTGCAGCTGAGACTTCCGGCTTCTCTT 120
Qy      101 SerLeuProGlnProSerIleProAlaAlaValProGlnSerAlaProProGluProHis 120
Db      121 TCCCTGCCCCCAGCCAGCATCCCCGGCTGTCCCGAGAGTGTCTCACCGGAGCCCAAC 180
Qy      121 ArgGluGluThrValThrAlaThrAlaThrSerGlnValAlaGlnGlnProProAlaAla 140
Db      181 CGGAGAGACCGGTACCCGACCGCACTTCCAGGTAGCCAGACGCTTCCAGCGCT 240
Qy      141 AlaAlaProGlyGluGlnAlaValAlaGlyProAlaProSerThrValProSerSerThr 160
Db      241 GCCGCCCTGGGGAAACAGCGCGTGGCGGCCCTGCCCCCTGCACGTCTCCCGACAGTACC 300
Qy      161 SerIysAspAlaProValSerGlnProSerLeuValGlySerIleGluGluProPro 180
Db      301 AGCAAAAGACCGCCAGTCTCCAGCTTGTGGGGAGCAAGAGAGAGCCGCCGCCG 360
Qy      181 AlaArgSerGlySerGlyGlyIleSerAlaIleGlyGluProGlnGluArgSerGlnGln 200
Db      361 GCGAAGATGGGAGCGCGCGCGAGCGCCAGAGGACCAAGAGAGAAAGAGAGAGAGAG 420
Qy      201 GlnAspAspIleGluGluLeuGluThrIleValAlaGlyMetSerAsnAspGlyArgPhe 220
Db      421 CAGATGATATCGAAGAGCTGGAAGCCAGGCGCGTGGGAATGTCTAAAGATGGCGCTT 480
Qy      221 LeuIlePheAspIleGluIleGlyArgGlySerPheIleThrValTyrIleGlyLeuAsp 240
Db      481 CTCAAGTTGACATCGAAATCGGACGAGGCTCTTTAAGAGCGTCTACAAAGTCTGAC 540
Qy      241 ThrGluThrThrValGluValAlaIleTyrCysGluLeuGlnAspArgIleThrIleSer 260
Db      541 ACTGAACACCGCGGAGAGCTGCTGCTGTGTAACGCGAGATCGAAATTAACAAAGTCT 600
Qy      261 GlnArgGlnArgPheIleGluGluAlaGluMetLeuIleGlyLeuGlnHisProAsnIle 280
Db      601 GAGAGGACGAGATTTAAAGAGAGCTGAATGTTTAAAGATCTTCAACATCCCAATTT 660
Qy      281 ValArgPheIleTyrAspSerTyrGluSerThrValIleGlyIleValIleValIleVal 300
Db      661 GTTAGATTATTTGATCTCTGGGAATCCACAGATTAAGAAAGAGATGATGTTTGGTG 720
Qy      301 ThrGluLeuMetThrSerGlyThrLeuIleThrIleIleIleIleIleIleIleIleIle 320
Db      721 ACTGAACCTTAGAGCTCTGGAACCTTAAACGATCTGAAAGAGTTTAAAGTGAAG 780
Qy      321 IleIleValIleAspSerTyrCysArgGlnIleLeuIleGlyLeuGlnPheLeuHisThr 340
Db      781 ATCAAGTTCTAAGAGAGTGGTGCCTGATCTTAAAGTCTTCAAGTTCTTCACTACT 840
Qy      341 ArgThrProProIleIleIleHisArgAspLeuIleCysAspAsnIlePheIleThrGlyPro 360
Db      841 CGAATCCACTATCATTCACCGGATCTTAAAGTGAACAACATCTTATCACCGGCCCT 900
Qy      361 ThrGlySerValIleIleGlyAspLeuGlyLeuAlaThrIleIleIleIleIleIleIle 380
Db      901 ACTGGCTCAGTCAGATTTGAGAGCTGCTGCTGGGACCTTGAAGCGGGCTTCTTTGCC 960
Qy      381 LysSerValIleGlyThrProGluPheMetAlaProGluMetTyrGluGluIleTyrAsp 400
Db      961 AAGAGTGTGATAGTACCCAGAGTTCATGCGCTGAGATGATGAGGAATAATGAT 1020
Qy      401 GluSerValAlaPheValTyrAlaPheGlyMetCysMetLeuGluMetAlaThrSerGluTyr 420
Db      1021 GAATCCGTTGACGTTTATGCTTTTGGAGTGTGATGAGTGTGATGAGGCTTCACTGATAT 1080
Qy      421 ProTyrSerGluCysGlnAsnAlaAlaGlnIleTyrArgArgValThrSerGlyValIle 440
Db      1081 CTTTACTCGAGTCCAAATGCTGCAAGATCTTACCGTCCGAGACCAAGTGGGGATGAG 1140
Qy      441 ProAlaSerPheAspIleValAlaIleProGluValIleGluIleIleGluIleCysIle 460

```

```
Db 1141 CCAGGCACTTTTGACAAAGTACCAATTCCTGAGAGGAAAGAAATTATTGAGAGTGCATA 1200
Qy 461 ArgGlnAsnLysAspGluArgTyrSerIleLysAspLeuLeuAsnIleAlaPheHesGln 480
Db 1201 CGACAAACAAAGAGTAAAGATATTCATCAAAAGACCTTTTGAAACCATGCTCTTCCAA 1260
Qy 481 GluGluThrGlyValArgValGluLeuAlaGluGluLysAspGlyGluLysIleAlaIle 500
Db 1261 GAGGAAACAGAGAGTACGGGTAGAAATTTAGCAGAAAGATGATGAGAAATAATATACCATTA 1320
Qy 501 LysLeuTyrLeuArgIleGluAspIleLysLysLeuLysGlyLysTyrLysAspAsnGlu 520
Db 1321 AATATTATGCTACGATTTAGAAATATTAAAGAAATTAAAGGAAATATACAAAGATTAATGA 1380
Qy 521 AlaIleGluPheSerPheAspLeuGluArgAspValProGluAspValAlaGlnGluMet 540
Db 1381 GCTATTGAGTTTCTTTGATTATTAAGAGAGATGTCCAGAAAGATGTTGCACAAAGAAATG 1440
Qy 541 ValGluSerGlyTyrValCysGluGlyAspHisLysThrMetAlaLysAlaIleLysAsp 560
Db 1441 GTAGAGTCTGGGTATGTCTGTAAAGGTATCAAAAGACCATGGCTTAAAGCTATCAAAAGAC 1500
Qy 561 ArgValSerLeuIleLysArgLysArgGluGlnArgGlnLeuValArgGluGluGlnGlu 580
Db 1501 AGAGTATCATTTATTAAGAGAAACGAGAGCGCGGAGTTGGTACGGAGAGAGCAAGAA 1560
Qy 581 LysLysLysGluGluGluSerSerLeuLysGlnValGluGlnSerSerAlaSerGln 600
Db 1561 AAAAAAAGCAGAGAGAGAGCGAGTCTCAACAGCAGGTAAAGAAATCAAGTCTTCCAG 1620
Qy 601 ThrGlyIleLysGluLeuProSerAlaSerThrGlyIleProThrAlaSerThrThrSer 620
Db 1621 ACGAGAAATCAAGCACTCTCTCTGTAGACCGGATACCTACGCTTCTTCAACCTTCA 1680
Qy 621 AlaSerValSerThrGlnValGluProGluGluProGluAlaAspGlnHisGlnGlnLeu 640
Db 1681 GCTTCAAGTTTCTCAACAAGTAGAACCTGAGAACTGAGGAGAGATCAACATCAACATCA 1740
Qy 641 GlnTyrGlnGlnProSerIleSerValLeuSerAspGlyThrValAspSerGlyGlnGly 660
Db 1741 CAGTCCACGACCAACCAATATCTGTATCTGATGAGACCGGTGACAGTGTCTCAGGGA 1800
Qy 661 SerSerValPheThrGluSerArgValSerSerGlnGlnThrValSerTyrGlySerGln 680
Db 1801 TCTCTGCTTCAAGAAATCTCGAGAGCAAGCAAGCAAGCTTCAATATGTTCCAA 1860
Qy 681 HisGlnGlnAlaHisSerThrGlyThrValProGlyHisIleProSerThrValGlnAla 700
Db 1861 CATGAACAGGCACTTCTACAGGACAGTCCAGGAGCATATACCTTCTACCTCCAGCA 1920
Qy 701 GlnSerGlnProHisGlyValTyrProProSerSerValAlaGlnGlyGlnSerGlnGly 720
Db 1921 CAGTCTCAGCCCATGGGTATATCCACCTCAAGTGTGGCACAAGGGGAGGCGCAAGGTT 1980
Qy 721 GlnProSerSerSerSerLeuThrGlyValSerSerSerGlnProIleGlnHisProGln 740
Db 1981 CAGCCATCTCAAGTAGTCTTACAGGGGTTTCACTTCCCAACCCATACAAATCTCCAG 2040
Qy 741 GlnGlnGlnGlyIleGlnGlnThrAlaProProGlnGlnThrValGlnTyrSerLeuSer 760
Db 2041 CAGCAGCAGGGAATACAGCAGACAGCCCTCTCAACAGACAGTGCAGATTCACCTTCA 2100
Qy 761 GlnThrSerThrSerSerGluAlaThrThrAlaGlnProValSerGlnProGlnAlaPro 780
Db 2101 CAGACATCAACCTCCAGTAGGCGCACTACTGCACAGCCAGTGCAGCTCAAGCTCAAGCTCA 2160
Qy 781 GlnValIleuProGlnValSerAlaGlyLysGln----- 791
Db 2161 CAAGTCTGCTCAAGTATCACTGGAACACAGCTTCCAGTTTCCAGCAGCTACCAACT 2220
Qy 791 ----- 791

Db 2221 ATCCAGGCGAAGCTCAGATCCAGTTGCGACACAAACCTCGGTGTTCAGTCCACTCT 2280
Qy 791 ----- 791
Db 2281 GGTGCTCATTTCTTCCAGTGGGAGCAGCGCTCCTACTCTCTTCTCCTCAGTACCCT 2340
Qy 791 ----- 791
Db 2341 GTCTCTCAGATTCCATATCAATCTCTCATGTGTCTACGGCTCAGACAGTTTCTCATCC 2400
Qy 791 ----- 791
Db 2401 CTTCCCATCAAGATGGCAGCTGGCATTACTCAGCTCTGTCTCAGCTGGCTTCATCTGCT 2460
Qy 791 ----- 791
Db 2461 ACAACAGCTGCGATCCCGGGGGTATCAACTGTGTCTTCTAGTCACTTCCAAACCTTCTG 2520
Qy 791 ----- 791
Db 2521 CAGCTGTGACTCAGCTGCCAAGTCAGGTTCACCCAAGCTTCTTCAACAGCAGTTGAG 2580
Qy 791 ----- 791
Db 2581 TCCATGGAAATACAGACTTAACCTTGGACAGAGCTGTGAGTTCCACTTCTCTGAGAT 2640
Qy 791 ----- 791
Db 2641 GTTCTGTACAGGGCTTCCACCTTGCAGTGCACACAGTACCAGAGATTCAAAATAT 2700
Qy 791 ----- 791
Db 2701 GCTCCTCTTCCAGTGTGCTTCTGTTCATCCATTTCTAGAGTCTTCCCTTCCCTCCAG 2760
Qy 791 ----- 791
Db 2761 CCGACAGAGTACTGGCTACACCTGGTACTTCCACAGTGTGACAGCTTATGTGAA 2820
Qy 791 ----- 791
Db 2821 TCAATCTTTTATGTTCTTATGGGTGTGTAGAGAGCAGTTCAAGTGTCCAGCCAGGA 2880
Qy 792 ----- 792
Db 2881 GGGAGTTTAGCAAGACCCCACTACATCTCCAGCAAGCAGTTTGGAGAGTCTCAG 2940
Qy 795 GlyValSerGlnValAlaProAlaGluProValAlaValAlaGlnProGlnAlaThrGln 814
Db 2941 GGAAGTCTCAGGTTGCTCTGTAGACAGTGCACATTCAGATGTGCTCAGGlnAGAT 3000
Qy 815 ProThrThrLeuAlaSerSerValAspSerAlaHisSerAspValAlaSerGlyMetSer 834
Db 3001 CCGACACTTGTGGCTTCTCTGTAGACAGTGCACATTCAGATGTGCTCAGGlnAGAT 3060
Qy 835 AspGlyAsnGluAsnValProSerSerSerGlyArgHisGlnGlyValThrThrLysArg 854
Db 3061 GATGCAATGAGAAAGTCCCATCTTCCAGTGAAGGAGTGAAGAAAGTCAAAACGAG 3120
Qy 855 HisTyrArgLysSerValArgSerArgSerArgHisGlnGlyValThrThrLysArg 874
Db 3121 CATTAACCAAAATCTGTAAAGAGTGTCTCTCAACATGAAAACTTCAAGCCCAAAATTA 3180
Qy 875 ArgIleLeuAsnValSerAsnLysGlyAspArgValAlaGluCysGlnLeuGluThrHis 894
Db 3181 AGAATTTTGAATGTTTAAATTAAGAGAGCCGAGTGAATGATTCATTAAGAGCTCAT 3240
Qy 895 AsnArgLysMetValThrPheLysPheAspLeuAspGlyAspAsnProGluGluIleAla 914
Db 3241 AATAGGAATATGCTTACATTCATTAATTTGACCTAGATGTGACAAACCCGAGAGATAGCA 3300
Qy 915 ThrIleMetValAsnAsnAspPheIleLeuAlaIleGluArgGluSerPheValAspGln 934
Db 3301 ACAATTAATGTGAACAATGACTTTTCTAGCAATATGAGAGAGTGTGTGTGTGATCA 3360
```

```
QY 935 ValArgGluIleIleGluIysAlaAspGluMetLeuSerGluAspValSerValGluPro 954
DB 3361 GGGCAGAAATATTGAAAAAGCTGATGAATGCTCAGTGAAGATGCTGATGGAGACA 3420
QY 955 GluGlyAspGlnGlyLeuGluSerLeuGlnGlyLysAspAspTyrGlyPheSerGlySer 974
DB 3421 GAGGGGTGATCAGAGGATGAGAGGTCTACAGAGAAAGATGACTATGGCTTTTCAGGTCT 3480
QY 975 GlnIleLeuGluGlyGluPheLysGlnProIleProAlaSerSerMetProGlnGlnIle 994
DB 3481 CAGAAATTCAGAGAGAGATTCACCAACCAATTCCTGGCTTCCTCAATGCGACAGACAAA 3540
QY 995 GlyIleProThrSerSerLeuThrGlnValValHisSerAlaGlyArgArgPheIleVal 1014
DB 3541 GGCAATTCCTACAGATTCCTTAATCAAGTGTTCATTCGGGGAAAGGGGTTTATAGTG 3600
QY 1015 SerProValProGluSerArgLeuArgGluSerLysValPheProSerGluIleThrAsp 1034
DB 3601 AGTCTGTGGCCAGAAAGCCGATTCAGAGATCAAAAGTTTCCCAAGTGAATTAACAGAT 3660
QY 1035 ThrValAlaAlaSerThrAlaGlnSerProGlyMetAsnLeuSerHisSerAlaSerSer 1054
DB 3661 AAGTTGCTGCTCTTACAGCTTCAGAGCCCTGGAAATGAATCTGTCTCACTTCATCATCC 3720
QY 1055 LeuSerLeuGlnGlnAlaPheSerGluLeuArgAlaGlnMetThrGluGlyProAsn 1074
DB 3721 CTTAGCTTACAAACAGGCCCTTTCTGAACTTAGACGTGCCCAATGACAGAAAGACCCAA 3780
QY 1075 ThrAlaProProAsnPheSerHisThrGlyProThrPheProValValProProPheLeu 1094
DB 3781 AAGAGACCTCCAAACTTATAGTATCAGAGACCAATTTCCAGTGTAGCTCTCTTCTTA 3840
QY 1095 SerSerIleAlaGlyValProThrThrAlaAlaAlaThrAlaProValProAlaThrSer 1114
DB 3841 AGTAGATTTGCTGGAGTCCCAACAGACAGACAGACAGACAGACAGCTCCCTGGACAAAGC 3900
QY 1115 SerProProAsnAspIleSerThrSerValIleGlnSerGluValThrValProThrGlu 1134
DB 3901 AGCCCTCCATGATGATTCATTCACATCAAGTATTCAGTGTGAGGTACAGTGGCCACTGA 3960
QY 1135 GluGlyIleAlaGlyValAlaThrSerThrGlyValValThrSerGlyLeuProIle 1154
DB 3961 GAGGGGATTTGCTGGAGTGGCCACAGACAGAGTGTAGTAACTTCAGTGTGCTCCCA 4020
QY 1155 ProProValSerGluSerProValLeuSerSerValValSerSerIleThrIleProAla 1174
DB 4021 CCACCTGTGTGAATTCACAGATCTTTCAGCGTATGTTCAAGTATCAACATATACCTCA 4080
QY 1175 ValValSerIleSerThrThrSerProSerLeuGlnValProThrSerThrSerGluIle 1194
DB 4081 GTTGTCTCAATATCTACTACATCCCGTCACTTCAAGTCCCAATCCACATCATGATGATC 4140
QY 1195 ValValSerSerThrAlaLeuThrProSerValThrValSerAlaThrSerAlaSerAla 1214
DB 4141 GTTGTCTTCTAGTACAGCACTGTATCTTCAGTAAACAGTTTCAGCAACTTCAGCTCTGCA 4200
QY 1215 GlyGlySerThrAlaThrProGlyProLysProProAlaValValSerGlnGlnAlaAla 1234
DB 4201 GGGGGCACTGCTGATCCCAAGCTCTTAAGCTCCAGCTGATGATTCACAGAGGACACA 4260
QY 1235 GlySerThrThrValGlyAlaThrLeuThrSerValSerThrThrThrSerPheProSer 1254
DB 4261 GGCAGCACTACGTGGAGCCACATTAACATCAAGTTTCTACACCACTTCAATCCCAAGC 4320
QY 1255 ThrAlaSerGlnLeuSerIleGlnLeuSerSerSerThrSerThrProThrIleuAlaGlu 1274
DB 4321 ACAGCTTACAGCTGTCCATTCAGCTTACAGCACTACTTCACTTCACTTACTTGAAGTGA 4380
QY 1275 ThrValValValSerAlaHisSerLeuAspLysThrSerHisSerSerThrThrGlyLeu 1294
DB 4381 ACCGTGTAGTATTAGCGCACCTCACTAATGATGACATCTCATAGAGTAACTGAGATTG 4440
QY 1295 AlaPheSerLeuSerAlaProSerSerSerSerSerProGlyAlaGlyValSerSerTyr 1314
DB 4441 GCTTCTCCTCCTCTGACACCATCTTCTCTTCTCTCTGAGAGAGATGCTACTATAT 4500
QY 1315 IleSerGlnProGlyGlyLeuHisProLeuValIleProSerValIleAlaSerThrPro 1334
DB 4501 ATTCTCAGCCGTGGGTGCTGATCTTGTGATTCATTCAGTATGATCTTACTCTCT 4560
QY 1335 IleLeuProGlnAlaAlaGlyProThrSerThrProLeuLeuProGlnValProSerIle 1354
DB 4561 ATCTTCCCAAGCAGAGAGCACTTCTTACACCTTTATTAACCAAGTATCTAGTATC 4620
QY 1355 ProProLeuValGlnProValAlaAsnValProAlaValGlnGlnThrLeuIleHisSer 1374
DB 4621 CCACCTTGTGATCAGCTCTTGTCCATATGTGCTGTGTATGACAGACACTAATTTATAGT 4680
QY 1375 GlnProGlnProAlaLeuLeuProAsnGlnProHisThrHisCysProGluValAspSer 1394
DB 4681 CAGCTCAACAGAGCTTGTCTTCCCAACAGCCCACTACTCATTTGCTGAAGTATGATCT 4740
QY 1395 AspThrGlnProLysAlaProGlyLysAspAspIleLysThrLeuGlnGluLysLeuArg 1414
DB 4741 GATACACAACCCAAAGCTCTGGAATGATGACATTAAGACTCTAAGAAAGAAAGCTGGCG 4800
QY 1415 SerLeuPheSerGluHisSerSerSerGlyAlaGlnHisAlaSerValSerLeuGlnThr 1434
DB 4801 TCTCTGTTCAAGTAAACAGCTCATCTGAGCTCAGCATGCTCTGTCTTCACTGAGAGCC 4860
QY 1435 SerLeuValIleGluSerThrValThrProGlyLysProThrThrAlaValAlaProSer 1454
DB 4861 TCACATGATCATAGAGAGCAGCTGACACACAGGACATCCCACTAGCTGTGACACCAAGC 4920
QY 1455 LysLeuLeuThrSerThrThrThrCysLeuProProThrAsnLeuProLeuGlyThr 1474
DB 4921 AAACCTCTGACTTCTACCAAGTACTGTGTACCAACCAATTTACACTAGAAACA 4980
QY 1475 ValAlaLeuProValThrProValValThrProGlyGlnValSerThrProValSerThr 1494
DB 4981 GTTGCTTTGCCAGTTACACCAAGTGTCACTGTGGCAAGTTTCAACCCAGTCAAGCAT 5040
QY 1495 ThrThrSerGlyValLysProGlyThrAlaProSerLysProProLeuThrLysAlaPro 1514
DB 5041 ACTACATCAGAGTGAAGACTGAACTGCTCCCTCCACAGCACTCTAATCAAGGCTCG 5100
QY 1515 ValLeuProValAlaGlyThrGluLeuProAlaGlyThrLeuProSerGluGlnLeuProPro 1534
DB 5101 GTGCTGCAGGGGTACTGAACCTTCCAGAGTACTTCAACAGGAGCACTGCACT 5160
QY 1535 PheProGlyProSerLeuThrGlnSerGlnGlnProLeuGluAspLeuAspAlaGlnLeu 1554
DB 5161 TTTCAGAGACTTCTCTAACCAGTCCAGCAACTCTAGAGATCTTGAATGCTCAATTC 5220
QY 1555 ArgArgThrLeuSerProGlu***IleThrValThrSerAlaValGlyProValSerMet 1574
DB 5221 AGAAGAACACTAATTCAGAGATATATCATCATGACTTCTGCGGTGTGCTGTGCTCATG 5280
QY 1575 AlaAlaProThrAlaIleThrGluAlaGlyThrGlnProGlnLysGlyValSerGlnVal 1594
DB 5281 GCGGCTCAACAGCATATCAAGAGCAAGACAGCTCGAAGGGGTTCCTCAACTTC 5340
QY 1595 LysGluGlyProValLeuAlaThrSerSerGlyAlaGlyValPheLysMetGlyArgPhe 1614
DB 5341 AAAGAGGCCCTGTCTTACAGAACTAGTTCAAGAGCTGTGTTTTAAAGATGGAGCATTT 5400
QY 1615 GlnValSerValAlaAlaAspGlyAlaGlnLysGluGlyLysAsnLysSerGluAspAla 1634
DB 5401 CAGGTTCTGTGTGAGCAGAGGGTGGCCAGAAAGAGGGTAAATAAGTCAAGAAAGATCA 5460
QY 1635 LysSerValHisPheGluSerSerThrSerGluSerSerValLeuSerSerSerPro 1654
DB 5461 AAGTGTGTCAATTTGAATTCAGCACTCAAGAGTCTCAAGTGTATCAAGTATGATGATCA 5520
QY 1655 GluSerThrLeuValLysProGluProAsnGlyIleThrIleProGlyIleSerSerAsp 1674
```



```

Db      241 GCGGCCCTGGGAAACAGCCGTCGGGGGCGCTCCCTCGACTGTCCCGAGTACC 300
Qy      161 SerIyAspArgProValSerGlnProSerLeuValGlySerIySgluGlnProPro 180
Db      301 AGCAAAAGACCGCCCACTGTCACAGCTTGTGGGAGCAAGAGAGCCGCGCG 360
Qy      181 AlaArgSerGlySerIyGlyGlySerAlaIySgluProGlnGlnIyArgSerGln 200
Db      361 GCGAGAGAGTGGCAGCGCGCGCGAGCCGCAAGAGCCACAGAGAGAAAGAGCCAG 420
Qy      201 GlnAspAspIleGlnGlnIyLeuGlnIyThrIySAlaValGlyMetSerAsnAspGlyArgp 220
Db      421 CAGGATGATATGAAGAGCTGAGACCAAGCGCGGATGTCTAAACATGCGCGCTTT 480
Qy      221 LeuIySAspAspIleGlnIleGlyArgGlySerPheIySerThrValIyIySgluIyLeuAsp 240
Db      481 CTCAGATTGACATCGAAATCCGCGAGAGGCTCTTTAAGACGGTCTACAAAGGCTTGGAC 540
Qy      241 ThrGlnThrThrValGlnValAlaIyPyrCysGlnLeuGlnAspArgIySLeuThrIySer 260
Db      541 ACTGAAACACCGCTGGAAGTCCCTGGTGTGAATCGAGAGATCGAAATTAACAAAGTCT 600
Qy      261 GlnArgGlnArgPheIySgluGlnAlaGlnMetLeuIySgluIyLeuGlnIySProAsnIle 280
Db      601 GAGAGCGAGAGATTTAAAGAGAGAGCTGAAATGTAAAGGCTTCAGCATCCCAATATT 660
Qy      281 ValArgPheIyAspSerThrIySgluSerThrValIySgluIySLeuIySValIyLeuAla 300
Db      661 GTTAAATTTATGATTCCTCGGAAATCCACAGTAAAGAGAGAGAGAGTCTTTTGGTG 720
Qy      301 ThrGlnLeuMetThrSerGlyThrLeuIySThrIyIyLeuIySArgPheIyValMetIyS 320
Db      721 ACTGAATCTTATGACCTCTGGAACACTTAAACGTATCGAAAGGTTTAAAGTATGATGAG 780
Qy      321 IleIySValIleuArgSerThrPyrCysArgGlnIleLeuIySgluIyLeuGlnPheIySThr 340
Db      781 ATCAAGATCTTAAGAGCTGGTGGCGCTGAGATCTTAAAGGCTTCAGTTCCTTACT 840
Qy      341 ArgThrProProIleIleHisArgAspLeuIySAspAsnIlePheIleThrGlyPro 360
Db      841 CGAATCCACCTATATTCATTCACCGCAATCTTAAATGACACATCTTATTCACCGGCT 900
Qy      361 ThrGlySerValIySleIyAspLeuGlyLeuAlaThrIleuIySArgAlaSerPheAla 380
Db      901 ACTGGCTCAGTCAGATTCAGAGACCTCGGCTGGACCCCTGAACCGGGCTCTTTTCC 960
Qy      381 LysSerValIleGlyThrProGlnPheMetAlaProGlnMetIyGlnGlnIySThrAsp 400
Db      961 AAGAGTGTGATGTACCCCAAGATTCATGCGCCCTGAGATGTATGAGAGAGAAATATGAT 1020
Qy      401 GlnSerValAspValIyThrAlaPheGlyMetCysMetLeuGlnMetAlaThrSerGlnIy 420
Db      1021 GAATCCGTTGACGTTTATGCTTTTGGAGTGCATGCTTGAGATGGCTTACATCTGAATAT 1080
Qy      421 ProIySerGlnCysGlnAsnAlaIleGlnIleIyThrArgValThrSerGlyValIyS 440
Db      1081 CTTTACTCGAGAGTGCMAAATGCTGCACAGATCTACCGTCCGTGACACAGTGGGTGAG 1140
Qy      441 ProAlaSerPheAspIyValAlaIleProGlnValIySgluIleIleGlnGlyCysIle 460
Db      1141 CCAGCGAGTTTGAACAAAGTACCAATCTCTGAAAGTGAAGAAATATTAAGAGATGACATA 1200
Qy      461 ArgGlnLeuIySAspGlnArgIySerIleIySAspLeuLeuAsnIleAlaPhePheGln 480
Db      1201 CGACAAACAAAGATGAAGATATTCATCAACAGCCCTTTGAACCAAGCTCTTCCAA 1260
Qy      481 GlnGlnThrGlyValArgValGlnLeuAlaGlnGlnIyAspAspGlyGlnIySleIleAlaIle 500
Db      1261 GAGGAAACAGAGAGTACGGGTGAAATTAAGACAGAAAGATGATGAGAAATAATGCCATA 1320
Qy      501 LysLeuThrPheuArgIleGlnAspIleIySlySLeuIySgluIySThrIySAspAsnGln 520

```

```

Db      1321 AAATTATGCTACGATTGAAGATATTAAGAAATTAAGGAAATAACAAAGATATGAA 1380
Qy      521 AlaIleGlnPheSerPheAspLeuGlnIyArgAspValProGlnAspValAlaGlnGlnMet 540
Db      1381 GCTATTGAGTTCCTTTGATTTAGAGAGAGTGTCCCAAGAAAGATGTGCACAAAGAAAG 1440
Qy      541 ValGlnSerGlyIyValIyCysGlnIyAspHisIySThrMetAlaIySAlaIleIySAsp 560
Db      1441 GTAGAGTCTGGGTATGCTGTGAAGTGTATCAAGACCATGCTTAAAGCTTCAAAAGAC 1500
Qy      561 ArgValSerLeuIleIySArgIySArgGlnIyArgGlnIyLeuValArgGlnGlnGlnIy 580
Db      1501 AGAGTATCATTAATTAAGAGAAACAGAGAGAGCGCGCATTTGTGTACGGAGAGCAAGA 1560
Qy      581 LysIySlySglnGlnGlnSerIySgluIySglnIyGlnIyGlnIySerIySAspGln 600
Db      1561 AAAAAAGACGAGAGAGAGAGCTTCAACAGCAGGTGAAGCAATCCAGTCTCCAG 1620
Qy      601 ThrGlyIleIySglnLeuProSerAlaSerThrGlyIleProThrAlaSerThrIySer 620
Db      1621 ACAGGATCAAGCAGCTCCCTTCTGTAGCAGCCGCAATCACTACTGCTTCAACACTTCA 1680
Qy      621 AlaSerValSerThrGlnValGlnProGlnIyProGlnAlaAspGlnIySglnIyLeu 640
Db      1681 GCTTCAGTTTCTACACAAGTGAACCTGAAGAACTGAGGACAGATCAACATCAACACTA 1740
Qy      641 GlnIyArgGlnIyProSerIleSerValLeuSerAspGlyIyThrValAspSerGlySglnIy 660
Db      1741 CAGTCCAGCAACCCAGATATCTGTGTATCTGAATGAGGACGTTGACAGTGTGAGGGA 1800
Qy      661 SerSerValPheThrGlnSerArgValSerSerGlnIyThrValSerIyGlySerGln 680
Db      1801 TCCTCTGCTTCACAGAACTCCGAGTACAGCCCAACAGACAGATTCAATATGTTCCCA 1860
Qy      681 HisGlnGlnAlaHisSerThrGlyIyThrValProGlyIySleProSerThrValGlnAla 700
Db      1861 CATGAACAGGACACTTCTACAGGACAGCTCCAGGCAATATACCTTCACTGTCCAAAGCA 1920
Qy      701 GlnSerGlnProHisGlnIyValIyProProSerSerValAlaGlnGlnIyGlnSerGlnIy 720
Db      1921 CAGTCCAGCCCAAGGGGTATATCCACCTCAAGTGTG----- 1959
Qy      721 GlnProSerSerSerLeuThrGlyValSerSerSerGlnProIleGlnHisProGln 740
Db      1959 ----- 1959
Qy      741 GlnGlnGlnGlyIleGlnGlnThrAlaProProGlnIyThrValGlnIySerIySer 760
Db      1960 ---CAGCAGGGAATACAGCAGACAGCCCTCCTCAACAGACAGTGCATTTCACTTCA 2016
Qy      761 GlnThrSerThrSerSerGlnAlaThrThrAlaGlnProValSerGlnProGlnAlaPro 780
Db      2017 CAGACATCACTCCAGTGAAGGCCACTACTCAGACAGCCAGTGAAGCTCAAGCTTCA 2076
Qy      781 GlnValLeuProGlnValSerAlaGlyIySglnSerThrGlnIyValSerGlnValAla 800
Db      2077 CAAGCTTGCTCAAGATACAGCTGGAAGAAACAGAGTACTCAGGAGTCTCCAGGTTCT 2136
Qy      801 ProAlaGlnProValAlaValAlaGlnProGlnAlaThrGlnProThrThrLeuAlaSer 820
Db      2137 CCTGACAGCCAGTTCAGTACAGCAGCCCAAGCTACCCAGCCGACCACTTGGCTTCC 2196
Qy      821 SerValAspSerAlaHisSerAspValAlaSerGlyMetSerAspGlyAsnGlnValAla 840
Db      2197 TCTGTAGCACTGCATTCAGATGTGCTTCAAGTATGATGATGCAATGAGAAAGCTC 2256
Qy      841 ProSerSerSerGlyArgHisGlnGlyIyArgThrThrIySArgHisIyArgIySerVal 860
Db      2257 CCATCTTCAGTGAAGCAGTGAAGAGAAAGAACTCAAAACGGCATTAACGAAATCTGTA 2316
Qy      861 ArgSerArgSerArgHisGlnIySThrSerArgProIySLeuArgIleLeuAsnValSer 880
Db      2317 AGAGTGTGCTCTGACATGAAGAAATTCACGCCCAAAATTAAGAAATTTGAATGTTTCA 2376

```


QY 881 AsnlysgIYAAspArgValaIGlucySGlnleuGIuThrHisAsnArglySmetValThr 900
Db 2377 AATAAAGGAGACCGAGTAGAATGTCAATTAAGACCTCAATAAGGAAATGGTTACA 2436
QY 901 PheIysPheAspLeuAspGIYAAspAsnProGIuGIuIleAlaThrIleMetValAsnAsn 920
Db 2437 TTCAATTTGACCTGATGTGTACAAACCCGAGAGATAGCAACATTTATGTGTAACAAT 2496
QY 921 AspPheIleLeuAlaIleGIuArgIuSerPheValAspGlnValArgIuIleIleGIu 940
Db 2497 GACTTATATTCTAGCAATAGAGAGAGATGCTGTTGTGATCAAGTCGAGAAATTTATTGA 2556
QY 941 LysAlaAspGIuMetLeuSerGIuAspValSerValGIuProGIuGIuIYAAspGlnIleu 960
Db 2557 AAAGCTGATGAATGCTCAGTAGAGATGTCAAGTGTGAACCAAGAGGTGATAGGGGATTG 2616
QY 961 GluSerLeuGIuGIuLYsAspAspTYrGIYpHeSerGIYSerGIuIleLYsLeuGIuGIu 980
Db 2617 GAGAGCTTACAGAAAGATGACTATGCTTTTCAAGTTCTCAAGAAATTTGAGAGAGAG 2676
QY 981 PheIysGIuProIleProAlaSerSerMetProGIuGIuIleGIYIleProThrsSer 1000
Db 2677 TTCAACAACCAATTCGCGCTCTTCATGCCACAGCAAAATGCGATTCTTCCAGTTCT 2736
QY 1001 LeuThrGIuValIleHisSerAlaGIYArgArgPheIleValSerProValProGIuSer 1020
Db 2737 TTAACTCAGTTGTTCTTCTGCGGAGAGCGGTTTATGTAGTCTGTGCGCAAGAAC 2796
QY 1021 ArgLeuArgIuSerLYsValPheProSerGIuIleThrAspThrValAlaIleSerThr 1040
Db 2797 CGATTACGAGAAATCAAAAGTTTCCCGAGTGAATAACAGATACAGTTGCTGCTTACA 2856
QY 1041 AlaGIuSerProGIYMetCAsnLeuSerHisSerAlaSerSerLeuSerLeuGIuGIuAla 1060
Db 2857 GCTCAGAGCCCTGGAATGAACTTGCTCACTGTCATCATCCCTTAGCTTACAAAGGCGC 2916
QY 1061 PheSerGIuLeuArgArgIleGIuMetThrGIuGIuProAsnThrAlaProProAsnPro 1080
Db 2917 TTTTCTGAACCTTAGAGTCCCAATAGACAAAGGCCCAATACAGCACCTTCAACCTTT 2976
QY 1081 SerHisThrGIYProThrPheProValValProProPheLeuSerSerIleAlaGIYVal 1100
Db 2977 ACTCATACAGAGCAACAACTTCCAGTAGTACCTCTTTTAAAGTACATGCTGAGTGC 3036
QY 1101 ProThrThrAlaAlaThrAlaProValProAlaThrSerSerProProAsnIle 1120
Db 3037 CCACCCACAGCAGCCACACACACAGCCCTGCAACAGCAGCCCTCTATGACAT 3096
QY 1121 SerThrSerValIleGIuSerGIuValThrValProThrGIuGIuGIYIleAlaGIYVal 1140
Db 3097 TCCACATCTCAATTCAGTCTAGGTTACAGTGCCACTGAAGAGGGGATTCCTGAGTT 3156
QY 1141 AlaThrSerThrGIYValValThrSerGIYGIYLeuProIleProProValSerGIuSer 1160
Db 3157 GCCACAGCAGCAGGTGTGTAACTTCAGGTGTCTCCCATACCACTGTGTCTGATCA 3216
QY 1161 ProValLeuSerSerValValSerSerIleThrIleProAlaValValSerIleSerThr 1180
Db 3217 CCAGACTTTCAGAGTAGTTCAGATACACAAATCCGAGGTGTCTCAATATCTACT 3276
QY 1181 ThrSerProSerLeuGIuValProThrSerThrSerGIuIleValValSerSerThrAla 1200
Db 3277 ACATCCCGCTCACTTCAATCCCGACATCACTGAGATGCTGTCTTCTGATACAGA 3336
QY 1201 LeuTYrProSerValThrValSerAlaThrSerAlaSerAlaGIYGIYSerThrAlaThr 1220
Db 3337 CTGTATCTTCTCAGTAAAGTTCAGCAACTTCAAGCTTGCAGGGGGAGTACTGCTAC 3396
QY 1221 ProGIYProLYsProProAlaValValSerGIuGIuAlaGIYSerThrThrValGIY 1240
Db 3397 CCAGGTCTTAAGCTTCAGCTGTAGTATCTCAGCAGGAGCAGCAGCACTACTGTGGGA 3456

QY 1241 AlaThrLeuThrSerValSerThrThrThrSerPheProSerThrAlaSerGIuLeuSer 1260
Db 3457 GCCACATTAACATCACTTCTTACCACTTCATTCCTCAAGCAGAGCTTCAAGCTGTCC 3516
QY 1261 IleGIuLeuSerSerSerThrSerThrProThrLeuAlaGIuThrValValSerAla 1280
Db 3517 ATTCAGCTTAGAGCAGTACTTCTACTCTCTTACTTACTGTAACCGGTGTAGTTCAGCG 3576
QY 1281 HisSerLeuAspLYsThrSerHisSerSerThrThrGIYLeuAlaPheSerLeuSerAla 1300
Db 3577 CACTCAGATGAATGAACATCTCATAGCAGTACACACTGGAATGGCTTCTCCCTCTCGA 3636
QY 1301 ProSerSerSerSerSerProGIYAlaGIYValSerSerTYrIleSerGIuProGIYGIY 1320
Db 3637 CCATCTTCTCTTCTCTCTCTGAGCAGAGAGTCTTATTTATTTCTCAGCTGTGGG 3696
QY 1321 LeuHisProLeuValIleProSerValIleAlaSerThrProIleLeuProGIuAlaAla 1340
Db 3697 CTGCATCTTGTGATTCATCAAGTAGAGCTTCACTCTTATCTTCCCAACAGACA 3756
QY 1341 GIYProThrSerThrProLeuProGIuValProSerIleProProLeuValGIuPro 1360
Db 3757 GAGCTACTTCTTACCTTTTATTTATCCCAAGTACCTAGATCCACCTTGTGTACGCT 3816
QY 1361 ValAlaAsnValProAlaValGIuGIuThrLeuIleHisSerGIuProGIuProAlaLeu 1380
Db 3817 GTTGCCATATGTCCTGCTGTACAGCAGACATTAATCATAGCAGCTCAACCACTTGG 3876
QY 1381 LeuProAsnGIuProHisThrHisCysProGIuValAspSerAspThrGlnProLYsAla 1400
Db 3877 CTTCACCAACAGCCCATCATCTATGTCTGTAAGTAGTTCTGATACCAACCCAAACT 3936
QY 1401 ProGIYIleAspAspIleLYsThrLeuGIuLYsLeuArgSerLeuPheSerGIuHis 1420
Db 3937 CCTGAATGTGAATGAATGAAGCTTGAAGAAAGGCTGGGTCTGTCTGTACGTAAAC 3996
QY 1421 SerSerSerGIYAlaGIuHisAlaSerValSerLeuGIuThrSerLeuValIleGIuSer 1440
Db 3997 AGCTATCTGAGTCCAGACATCCCTGTCTGTACCTGAGACCTCACTAGTCAATAGAGC 4056
QY 1441 ThrValThrProGIYIleProThrThrAlaValAlaProSerLYsLeuLeuThr 1460
Db 4057 ACTGCACACAGCAGGATCCCACTACTGCTGTGACCAAGCAAACTCCTGACTTCAAC 4116
QY 1461 ThrSerThrCysLeuProProThrAsnLeuProLeuGIYThrValAlaLeuProValThr 1480
Db 4117 ACAGATACTTGTCTTACCAACCAATTTACCATAGGAACGTGTGTGCTGCAAGTTACA 4176
QY 1481 ProValValThrProGIYGIuValSerThrProValSerThrThrSerGIYValLYs 1500
Db 4177 CCAGTGTGACACCTGGGCAAGTTCTTACCCAGTCAACAGTCACTACATCAGAGTGA 4236
QY 1501 ProGIYThrAlaProSerLYsProProLeuThrLYsAlaProValLeuProValGIYThr 1520
Db 4237 CTGGAAGCTGCTCCCTCCAGGACACTCTAACTAAGGCTCGGCTGCTGCACTGTGACT 4296
QY 1521 GIuLeuProAlaGIYThrLeuProSerGIuGIuLeuProProPheProGIYProSerLeu 1540
Db 4297 GAAGTTCAGAGGAGTACTTACCGCAGCAGAGCTGCCACTTCTTCCAGAACCTTCTCTA 4356
QY 1541 ThrGIuSerGIuGIuProLeuGIuAspLeuAspAlaGIuLeuArgArgThrLeuSerPro 1560
Db 4357 ACCCAGTCCACAGCAACCTTACAGATCTTGATGTCAATTAAGAAACACTTGTGCTCA 4416
QY 1561 GIu***IleThrValThrSerAlaValGIYProValSerMetAlaAlaProThrAlaIle 1580
Db 4417 GAGATATCTACAGTACTTCTGCGGTGTGCTGTGCTCATGCGGCTCAACACAAATC 4476
QY 1581 ThrGIuAlaGIYThrGIuProGIuLYsGIYValSerGIuValLYsGIuGIYProValLeu 1600
Db 4477 ACAGAAAGCAGGAACAGCTTCAAGAGGAGTGTCTCAAGTCAAAAGAGCCCTGTCTCA 4536
QY 1601 AlaThrSerSerGIYAlaGIYValPheLYsMetGIYArgPheGIuValSerValAlaAla 1620


```

Db      4537 GGAACCTAGTTCAGAGAGCTGGTGTTTTAAGATGGAGCAATTCAGGTTCTTCTTCAGCA 4596
Qy      1621 AaSPglYAlaGlnlySGlyUGlyLysAsnlySSerGIuAspAlaLysSerValHisphEgIU 1640
Db      4597 GACGGGTGCCCAAGAAAGAGGTAAAAATTAAGTCAAGAAAGATGCAAAAGTCTTCAATTTTGA 4656
Qy      1641 SerSerThrSerGIuSerSerValLeuSerSerSerSerProGIuSerThrLeuValLys 1660
Db      4657 TCCAGACACCTCAGAGTCTCAGTGTCTATCAAGTAGTCCAGAGATGACCTTGGTGA 4716
Qy      1661 ProGIuProAsnGlyIleThrIleProGlyIleSerSerAspValProGIuSerAlaHis 1680
Db      4717 CCAGAGCCGAATGGATACCACTCCCTGTATCTTTCAGATGTCCAGAGAGTGCCAC 4776
Qy      1681 LysThrThrAlaSerGIuAlaLysSerAspThrGIuGIuProThrLysValGIuArgPhe 1700
Db      4777 AAAACTACTGCTTCAGAGGCAAGTACGACCTGGGCGACCTAACCAAGGTTGGACGTTT 4836
Qy      1701 GlnValIleThrThrAlaAsnLysValGIuArgPheSerValSerLysThrGIuAspLys 1720
Db      4837 CAGGTGACAACTACAGCAAAACAAGTGGGTGTTCTGTATCAAAAACAGAGACAAAG 4896
Qy      1721 IleThrAspThrLysLysGIuGIuProValAlaSerProProPheMetAspLeuGIuGln 1740
Db      4897 ATCACTGACACAAAGAAAGAGACAGTGGCATCTCTCTTTATGGATTGGACAA 4956
Qy      1741 AlaValLeuProAlaValIleProLysLysGIuLysProGIuLeuSerGIuProSerHis 1760
Db      4957 GGTGTCTCTCTCGTGTATACCAAGAAAGAAAGCCGAACTGTCCAGAGCTTCACAT 5016
Qy      1761 LeuAsnGIuProSerSerAspProGIuAlaAlaPheLeuSerArgAspValAspAspGIu 1780
Db      5017 CTAATGGGCGCTCTTCTGACCCGAGGCGCTTTTAAAGTAGGAATGGATGATG 5076
Qy      1781 SerGIuSerProHisSerProHisGIuLeuSerSerLysLeuProSerGIuAsnLeu 1800
Db      5077 TCCGGTATGTCCACACTCCGCCATCAGCTCAAGACCTCAAGAGCTTCACAGATCTA 5136
Qy      1801 SerGlnSerLeuSerAsnSerPheAsnSerSerIleMetSerSerAspAsnGIuSerAsp 1820
Db      5137 AGTCAAAAGCTTAAATTCATTATTCATCTCTTTCATGATGAGGCAAAAGAGTCAAGT 5196
Qy      1821 IleGIuAspGIuAspLeuLysLeuGIuLeuArgArgLeuArgAspLysHisLeuLysGIu 1840
Db      5197 ATCGAAGATGAAGACTTAAGTAGAGCTGCGACGACACTACGATAAACATCTCAAGAG 5256
Qy      1841 IleGlnAspLeuGlnSerArgGlnLysHisGIuIleGIuSerLeuThrThrLysLeuGIu 1860
Db      5257 ATTCAGACCTGCAGAGTCCGCGAAGCATGMAATTGATCTTTGTATACCAAACTGGGC 5316
Qy      1861 LysValProProAlaValIleIleProProAlaAlaProLeuSerGIuArgArgArg 1880
Db      5317 AAGGTGCCCCCTGCTGTATTAATTCCTCCCGAGCTGCTCCCTTCAGGAGAAAGACAGCA 5376
Qy      1881 ProThrLysSerLysGIuSerLysSerSerArgSerSerSerLeuGIuAsnLysSerPro 1900
Db      5377 CCCACTAAAGCAAGGACGCAAAATCTAGTGAAGCACTTCTTGGGAAATAAAGCCCC 5436
Qy      1901 GlnLeuSerGIuAsnLeuSerGIuGlnSerAlaAlaSerValLeuHisProGIuGlnThr 1920
Db      5437 CAGCTTTAGGTAACTCTGTGTGTGTCAGAGTGCAGCTTGCAGCTTGCACCCACAGAC 5496
Qy      1921 LeuHisProProGIuAsnIleProGIuSerGIuGlnAsnGIuLeuLeuGIuProLeuLys 1940
Db      5497 CTCACCTCTCTGGCAATCTCCAGAGTCCGGGCGAAGTACGCTTTTCAGCCCTTTAAG 5556
Qy      1941 ProSerProSerSerAspAsnLeuThrSerAlaPheThrSerAspGIuAlaIleSerVal 1960
Db      5557 CCATCTCCCTCCAGAGCAACCTTATTCAGCTTTCACAGTGAAGTGGCCATTTCACTA 5616
Qy      1961 ProSerLeuSerAlaProGIuGlnGIuThrSerSerThrAsnThrValGIuAlaThrVal 1980

```

```

Db      5617 CCAGACCTTTCTGCTCCAGTCAAGAAACAGACACACAAACTGTTGGGCGCAACAGTG 5676
Qy      1981 AenSerGIuAlaGlnAlaGlnProProAlaMetThrSerSerArgLysGIuThrPhe 2000
Db      5677 AACAGCCAAAGCCGCAAGCTCAGCTCTCTGCAAGACGTCCAGACGAAAGGCGCACTTC 5736
Qy      2001 ThrAspLeuHisLysLeuValAspAsnThrAlaArgAspAlaMetAsnLeuSerGIu 2020
Db      5737 ACAGATGACTTGCACAAATTTGGTAGCAATTTGGGCCGAGATGCCATGAATCTTCCAGGC 5796
Qy      2021 ArgArgLysSerLysGIuHisMetAsnThrGIuGIuProGIuMetAlaArgLysPheSer 2040
Db      5797 AGGAGAGGAAACAAAGGCGCACATGATATGAGGGCCCTGGAATGGCAAGAAATTTCTCT 5856
Qy      2041 AlaProGIuGlnLeuCysLleSerMetThrSerAsnLeuGIuLysSerAlaProIleSer 2060
Db      5857 GCACCTGGGCAACTGTGATCTCCATGACCTTCGAACTCGGGGTGGCTCCGCCCATCTCT 5916
Qy      2061 AlaAlaSerAlaThrSerLeuGIuHisPheThrLysSerMetCysProProGIuGlnThr 2080
Db      5917 GCAGCATGACTACTCTCTAGTCACTTCCACAAATGATGTGCCCCCAGACAGATAT 5976
Qy      2081 GlyPheProAlaThrProPheGIuAlaGlnThrSerGIuThrGIuProAlaProGln 2100
Db      5977 GGCCTTCCAGCTACCCCATTTGGCGCTCAATGAGTGGAGCGGGTGCCAGCACACAG 6036
Qy      2101 ProLeuGIuGlnPheGlnProValGIuThrAlaSerLeuGlnAsnPheAsnIleSer 2120
Db      6037 CCACCTGGCCAGTTCCAACTGTGGAACTGCTCTTTCAGAAATTTCAACATCCAGCAAT 6096
Qy      2121 LeuGlnLysSerLleSerAsnProProGIuSerAsnLeuArgThrThr 2136
Db      6097 TTGCAAGAAATCCATCAGCAACCCCGACGCTCCAACTCGGAGCACACT 6144

RESULT 14
AR302615 6426 bp DNA linear PAT 12-JUN-2003
LOCUS AR302615
DEFINITION Sequence 55 from patent US 6541252.
ACCESSION AR302615
VERSION AR302615.1 GI:31690913
KEYWORDS
SOURCE
ORGANISM
Unknown.
REFERENCE
1 (bases 1 to 6426)
AUTHORS
Walke, D.W., Hilbun, F., Donoho, G. and Turner, C.A. Jr.
TITLE
Human kinases and polynucleotides encoding the same
JOURNAL
Patent: US 6541252-A 55 01-Apr-2003;
FEATURES
Location/Qualifiers
1..6426
/molecule="unknown"
/organism="human"
/mol_type="genomic DNA"

ORIGIN
Alignment Scores:
Pred. No.: 4,12e-245 Length: 6426
Score: 10287.50 Matches: 2048
Percent Similarity: 94.42% Conservative: 0
Best Local Similarity: 94.42% Mismatches: 0
Query Match: 95.15% Indels: 121
DB: Gaps: 2

US-10-010-720-14 (1-2136) x AR302615 (1-6426)
Qy      61 MetAspLysAspSerAlaArgGIuAlaAlaThrThrThrThrGlnHisArgPhePhe 80
Db      1 ATGACAAAGACACACCTGGGGCGCGCGCGACCACTTCCACACCTGAGCACCGCTTCTTC 60
Qy      81 ArgArgSerValIleCysAspSerAsnAlaThrAlaLeuGlnLeuProGIuLysProLeu 100
Db      61 CGCGAGAGCGTCACTGTGAGTCCAAATGCCATGCACTGGAGCTTCCGGGCTTCTCTT 120
Qy      101 SerLeuProGlnProSerIleProAlaAlaValProGlnSerAlaProProGlnProHis 120

```

```

Db      121  TCCCTCCGCCAGCCAGCATCCCGGCTGTCCGCAAGTGTCCACCGGAGGCCAC 180
Qy      121  ArgGluGluThrValThrAlaThrSerGlnValAlaGlnGlnProProAlaAla 140
Db      181  CGGGAAAGAGACCGGTACCGCCACCCCTCCAGTAGGCCACAGCCTCCACCGCT 240
Qy      141  AlaAlaProGluGluGlnAlaValAlaGlyProAlaProSerThrValProSerSerThr 160
Db      241  GCCGCCCTGGGGAAACAGGCGCTCGGGGCTCCCTGCGCTCCGACGTGTCCGCAAGTACC 300
Qy      161  SerIysAspArgProValSerGlnProSerLeuValGlySerIysGluGluProProPro 180
Db      301  AGCAAAAGACCGCCAGCTGTCCAGCCTAGCTGTGGGGAGCAAAAGAGAGCGCGCCG 360
Qy      181  AlaArgSerGlySerGlyGlyGlySerAlaIysGluProGlnGluGluArgSerGln 200
Db      361  GCGAAGAGTGGCGAGCGCGCGCGCGCAAGAGCAAGAGAGAAAGAGAGCAAGAG 420
Qy      201  GlnAspAspIleGluGluLeuGluThrIysAlaValAlaGlyMetSerAspAspGlyArgPhe 220
Db      421  CAGGATGATATGAAAGAGCTGGAGAACCAAGCGCGGAGATGTCTAACGATGCGCGCTT 480
Qy      221  LeuIysPheAspIleGluIleGlyArgGlySerPheIysThrValTyrIysGlyLeuAsp 240
Db      481  CTCAAGTTTGACATGAAATCGGCAAGGCTCTTTAAGACGCTTACAAAGCTCTGAGC 540
Qy      241  ThrGluThrThrValGluValAlaTyrCysGluLeuGlnAspArgIysLeuThrIysSer 260
Db      541  ACTGAACCAACCGCTGAGAGTCCCGTGGTGTGAATCGAAGATCGAAATTAACAAAGCT 600
Qy      261  GlnArgGlnArgPheIysGluGluAlaGluMetLeuIysGlyLeuGlnHisProAsnIle 280
Db      601  GAGAGCGAGAGATTTAAAGAAAGAGCTGAATGTTAAAGTCTTCAAGATCCCAATATT 660
Qy      281  ValArgPheIysAspSerTyrIysLeuSerThrValIysGlyIysCysIleValLeuVal 300
Db      661  GTTAATTTTATGATTCCTCGGATCCACAGTAAAGAAAGAAAGAGATGCTTTTGGTG 720
Qy      301  ThrGluLeuMetThrSerGlyThrLeuIysThrTyrLeuIysArgPheIysValIysLeu 320
Db      721  ACTGAACCTTATACGCTGTGGAACACTTAAACGTAATCGAAAGGTTTAACTGATGAAG 780
Qy      321  IleIysValIleuArgSerTyrCysArgGlnIleLeuIysGlyLeuGlnPheLeuHisIle 340
Db      781  ATCAAGTTCTTAAGAGCTGGTGGCTCGATCTTAAGTCTTCAAGTTCTTCACTACT 840
Qy      341  ArgThrProProIleIleHisArgAspLeuIysCysAspAsnIlePheIleThrGlyPro 360
Db      841  CGAACTCCACCTATCATTCACCGCATCTTAATGTGACAAACATCTTATCACCGGCGCT 900
Qy      361  ThrGlySerValIysIleGlyAspLeuGlyLeuAlaThrLeuIysArgAlaSerPheAla 380
Db      901  ACTGGCTCAGTAAAGATTGGAGACCTCGGTCTGGCAACCTGAACCGGGCTCTTTTGGC 960
Qy      381  LysSerValIleGlyThrProGlnPheMetAlaProGluMetTyrGluIysTyrAsp 400
Db      961  AAGAGTGTATAGTATGCCCAAGTTCATGCGCCCTCGAGATGTATGAGAGAAATGTGAT 1020
Qy      401  GlnSerValAspValTyrAlaPheGlyMetCysMetLeuGluMetAlaThrSerGluTyr 420
Db      1021  GAATCCGTTGACGTTTATGCTTTTGGGATGTGATGCTTGAAGTGGCTCATCTGAAT 1080
Qy      421  ProTyrSerGluCysGlnAsnAlaAlaGlnIleTyrArgArgValThrSerGlyValIys 440
Db      1081  CCTTACTGGAGTGGCCAAATGTCTGCACAGATCTACCGTGGTGAACAGTGGGGTGAAG 1140
Qy      441  ProAlaSerPheAspIysValAlaIleProGluValIysGluIleIleGluGlyCysIle 460
Db      1141  CCAAGCAAGTTTGAACAAATAGCAATTCCTGAAGTGAAGAAATTTTGAAGAGATGATA 1200
Qy      461  ArgGlnAsnIysAspGluArgTyrSerIleIysAspLeuLeuAsnHisAlaPhePheGln 480

Db      1201  CGAACAACAAAGATGAAGATATTCATCAAGACCTTTTGAACCATGCTTTCCAA 1260
Qy      481  GluGluThrGlyValArgValGluLeuAlaGluGluAspArgGlyGluIysIleAlaIle 500
Db      1261  GAGGAAACAGAGATACGGGTGAATTAACAGAAAGATGATGAGAAATTAATACCAATA 1320
Qy      501  LysLeuTyrLeuAspIleGlyAspIleIysIysLeuIysGlyIysTyrIysAspAsnIle 520
Db      1321  AAATTTATGGCTATGATTAAGAAATTTAAAGAAATTTAAAGAAATTTAAAGATTAATGA 1380
Qy      521  AlaIleGluPheSerPheAspLeuGluArgAspValProGluAspValAlaGlnIleMet 540
Db      1381  GCTATTAGATTTTCTTTGATTTAGAGAGATGTCCCAAGAAAGATTTGCACAAAGAAATG 1440
Qy      541  ValGluSerGlyTyrValCysGluGlyAspHisIysThrMetAlaIleAlaIleIysAsp 560
Db      1441  GTAGAGTGGGTATGTGTGTGAAGTATCAAGAGACATGCTTAAGCTATCAAGAGC 1500
Qy      561  ArgValSerLeuIleIysArgIysArgGluGlnArgGlnLeuValArgGluGluGlnIle 580
Db      1501  AGAGTATCATTAATTAAGAGAAACGAGACGCGGAGTTGTGTCGGAGAGAGCAAGA 1560
Qy      581  LysIysIysGlnGluGluSerSerLeuIysGlnIleValGluGlnSerSerAlaSerGln 600
Db      1561  AAAAAAAGCAGAAAGAGAGAGCTCAACAGCAGTAGAACAATCCAGTGTCTCCAG 1620
Qy      601  ThrGlyIleIysGlnLeuProSerAlaSerThrGlyIleProThrAlaSerThrIysSer 620
Db      1621  ACAGGAATCAAGACAGCTCTCTTGTCTAGACCGGCAATCACTACTGTCTTCAACCTTCA 1680
Qy      621  AlaSerValSerThrGlnValGluProGluGluProGluAlaAspGlnHisGlnGlnLeu 640
Db      1681  GCTTCAGTTTCTACCAAGTAGAAGACTCAAGACCTGAGGCAAGATCAACATCAACACTA 1740
Qy      641  GlnTyrGlnGlnProSerIleSerValLeuSerAspGlyThrValAspSerGlyGlnGly 660
Db      1741  CAGTACCAAGAACCCAGATATCTGTATCTGATGAGGAGCGGTGAACAGTGTCAAGGA 1800
Qy      661  SerSerValPheThrGluSerArgValSerSerGlnGlnThrValSerTyrGlySerGln 680
Db      1801  TCCTTGCTTCAAGAAATCTGAGTGAAGCAAGCAAGCAAGTTCATTAATGATGTTCCAA 1860
Qy      681  HisGlnGlnAlaHisSerThrGlyThrValProGlyHisIleProSerThrValGlnAla 700
Db      1861  CATGAACGAGCAATTCACAGGCAAGATCCAGAGGATATACCTTCACTACTCCAAAGA 1920
Qy      701  GlnSerGlnProHisGlyValTyrProProSerSerValAlaGlnGlyGlnSerGlnGly 720
Db      1921  CAGTCTCAGCCCATGGGGTATATCCACCCTCAAGTGTG----- 1959
Qy      721  GlnProSerSerSerSerLeuThrGlyValSerSerSerGlnProIleGlnHisProGln 740
Db      1959 ----- 1959
Qy      741  GlnGlnGlnGlyIleGlnGlnThrAlaProProGlnGlnThrValGlnTyrSerLeuSer 760
Db      1960 ---CAGCAGGGAATACAGAGACAGCCCTCTCAACAGACAGTCAAGTATCACTTCA 2016
Qy      761  GlnThrSerThrSerSerGluAlaThrThrAlaGlnProValSerGlnProGlnAlaPro 780
Db      2017  CAGACATCAACCTCCAGAGAGGCCACTACTGCACGCCAGTGAAGTCAAGCTCAACCTCCA 2076
Qy      781  GlnValLeuProGlnValSerAlaGlyIysGln----- 791
Db      2077  CAAGTCTGCTCAAGATTCAGCTGAAAACAGGGCTTCCCACTCGACTGCACACAG 2136
Qy      791 ----- 791
Db      2137  TACCCAGAGATTAATATATGCTCCCTCTTCCAAAGTGGCTTGTGTCATCATTCCT 2196
Qy      791 ----- 791
Db      2197  ACAGTCTTACCCCTCCATGCCAGACAGAAATGACTGCTTACACTGGTACTTCCACA 2256

```

Qy	791	-----	791	Qy	1108	AlaProValProAlaThrSerSerProProAsnAspIleSerThrSerValIleGlnSer	1127
Db	2257	GTGTGCAGCCTTATGTGATCAAAATCTTTAGTTCCTATGGGTGTGTAGAGAGACAG	2316	Db	3337	GCACAGTCCCTGGACAAGACGCCCTCTATGACATTTCCACATCAGTATTCAGTCT	3396
Qy	791	-----	791	Qy	1128	GluValThrValProThrGluGluGlyIleAlaGlyValAlaThrSerThrGlyValVal	1147
Db	2317	GTTCAAGTGTCCAGCCAGAGGAGGAGTTTATGACAAAGCCCACTACCTCCAGCA	2376	Db	3397	GAGTTTACAGTCCCACTGAAGAGGGGATTCGTGAGTTGCCACAGCAGCAGGTGTGTA	3456
Qy	792	-----	792	Qy	1148	ThrSerGlyIleuProIleProProValSerGluSerProValLeuSerSerValVal	1167
Db	2377	GCAAGTTTGGAGAGTACTCAGGAGAGTCTCAGTTGCTCTCCGACAGCAGTTGCACTA	2436	Db	3457	ACTTGAAGTGTCTCCATACACCTGTGTGATGATACCAAGTACTTCCAGCGGATGT	3516
Qy	808	AlaGlnProGlnAlaThrGlnProThrThrLeuAlaSerSerValAspSerAlaHisSer	827	Qy	1168	SerSerIleThrIleProAlaValAlaSerIleSerThrThrSerProSerLeuGlnVal	1187
Db	2437	GCACAGCCCCCAAGTACCCAGCCAGCACTTGGCTTCTCTGTAGACGTGCACATTC	2496	Db	3517	TCAAGTATCAGATATCTCGAGTGTCTCATATTTACTACATCCCGCTCACTTCAAGTC	3576
Qy	828	AspValAlaSerGlyMetSerAspGlyAsnGluAsnValProSerSerSerGlyArgHis	847	Qy	1188	ProThrSerThrSerGluIleValValSerSerThrAlaLeuThrProSerValThrVal	1207
Db	2497	GATGTGCTTCAGATGATGATGATGCAATGAGAAAGTCCCATCTTCCAGTGAAGGCA	2556	Db	3577	CCCATCTCCATCATGAGATGCTGTGTTCTGTAGTACAGCAGCTGTATCTTCAGTAA	3636
Qy	848	GluGlyArgThrThrLysArgHisArgArgLysSerValArgSerArgSerArgHis	867	Qy	1208	SerAlaThrSerAlaSerAlaGlySerThrThrValGlyAlaThrLeuThrSerValSer	1227
Db	2557	GAAAGAAAGACTACAAAGCGCTTACCGAAATCTGTAAGAGTCTCTGACATGGA	2616	Db	3637	TCAGCACTTACAGCTCTGCAGGGGGAGTACTGTGTACCCAGGTCTTAAGCTTCAGCT	3696
Qy	868	LysThrSerArgProLysLeuArgIleLeuAsnValSerAsnLysGlyAspArgValVal	887	Qy	1228	ValValSerGlnGlnAlaAlaGlySerThrThrValGlyAlaThrLeuThrSerValSer	1247
Db	2617	AAATCTCACGCCCAAAATTAGAAATTTGAATGTTTCAATTAAGAGACCGACTAGTA	2676	Db	3697	GTAATATCTCAGCAGCAGCAGCAGCAGCTTCAGCTGCATTTAGCATTAGCAGCAGT	3756
Qy	888	GluGlyGlnLeuGluThrHisAsnArgLysMetValThrPheLysPheAspLeuAspGly	907	Qy	1248	ThrThrThrSerPheProSerThrAlaSerGlnLeuSerIleGlnLeuSerSerThr	1267
Db	2677	GAATGTCAATTAGAACCTCATATAGAAATGTTGATTAATCAATTAAGTCACTAGT	2736	Db	3757	ACCACACTTCAATCCCAAGCAGCAGCTTCACAGCTGCATTTAGCATTAGCAGCAGT	3816
Qy	908	AspAsnProGluGluIleAlaThrIleMetValAsnAspPheIleLeuAlaIleGlu	927	Qy	1268	SerThrProThrIleAlaGluThrValValSerAlaHisSerLeuAspLysThrSer	1287
Db	2737	GACAAACCCCGAGAGATAGCAACATTTGTGACAAATGACTTATTTCTGCAATTAAG	2796	Db	3817	TCATCTCTTACTTAACTGGAACCGGTGTACTGTAGCCGACACTCATGATTAAGCATCT	3876
Qy	928	ArgGlnSerPheValAspGlnValArgGluIleIleGluLysAlaAspGluMetLeuSer	947	Qy	1288	HisSerSerThrThrGlyLeuAlaPheSerLeuSerAlaProSerSerSerSerPro	1307
Db	2797	AGAGAGTGTGTTGATGATCAAGTCCGAGAAATTTGAAAAGCTGAGTGAATGTCT	2856	Db	3877	CATGACGATACATGATGATGCTTTCCTCTCTGACATCTCTCTCTCTCTCTCT	3936
Qy	948	GluAspValSerValGluProGluGlyAspGlnLysLeuGluSerLeuGlnGlyLysAsp	967	Qy	1308	GlyAlaGlyValSerSerTyrIleSerGlnProGlyGlyLeuHisProLeuValIlePro	1327
Db	2857	GAGGATGTCAAGTGGAAACGAGAGGTGATCAGGGATTTGAGAGCTCTCAAGGAAAGAT	2916	Db	3937	GGAGCAGAGAGTGTACTGATTAATTTCTCAGCTGTGGGCTGCATCTTGGTCA	3996
Qy	968	AspTyrGlyPheSerGlySerGlnLysLeuGluGlyGluPheLysGlnProIleProAla	987	Qy	1328	SerValIleAlaSerThrProIleLeuProGlnAlaAlaGlyProThrSerThrProLeu	1347
Db	2917	GACTATGCTTTTTCAGTTCTCAGAAATTTGAGAGAGAGTTCAAACCAATTCCTCG	2976	Db	3997	TCAGTATGATCTTACTCTCTATCTTCCCAACGACGAGGACCTACTTCTACACCTTTA	4056
Qy	988	SerSerMetProGlnGlnIleGlyIleProThrSerSerLeuThrGlnValAlaHisSer	1007	Qy	1348	LeuProGlnValProSerIleProProLeuValGlnProValAlaAsnValProAlaVal	1367
Db	2977	TCTTCATGCCACAGCAATAGGCAATCTCTACCAAGTCTTTAATCAAGTGTCTCAT	3036	Db	4057	TTACCCCAAGTACTAGTATCCACCTTGTGTACAGCTGTGGCAATGTGCTGCTGTA	4116
Qy	1008	AlaGlyArgArgPheIleValSerProValProGluSerArgLeuArgGluSerLysVal	1027	Qy	1368	GlnGlnThrLeuIleHisSerGlnProGlnProAlaLeuLeuProAsnGlnProHisThr	1387
Db	3037	GCGGGAAGCGGTTTATAGTAGTCTGTGCGAGAAAGCCGATTAACGAAATCAAAAGTT	3096	Db	4117	CAGCAGACACTAATTAATGATGACAGCTCAACCACTTGTCTCCCAACAGCCCACT	4176
Qy	1028	PheProSerGluIleThrAspThrValAlaAlaSerThrAlaGlnSerProGlyMetAsn	1047	Qy	1388	HisCysProGluValAspSerAspThrGlnProLysAlaProGlyIleAspAspIleLys	1407
Db	3097	TTCCCGATGAAATACAGATACAGTCTCTGCTCTACAGCTCAGAGCCCTCGAATGAC	3156	Db	4177	CATTGCTCTGAAGTAGATCTGATACACAAACCCAAAGCTCTGGAATTAAGACATAAG	4236
Qy	1048	LeuSerHisSerAlaSerSerLeuSerLeuGlnAlaPheSerGluLeuArgArgAla	1067	Qy	1408	ThrLeuGluGluLysLeuArgSerLeuPheSerGluHisSerSerSerGlyAlaGlnHis	1427
Db	3157	TTGTCTCACTCTGATCATCTCCCTTAGTCTCAACAGGCTTTTTCGAATTAAGAGTCC	3216	Db	4237	ACTCTAGAGAAAGCTGCGGTCTCTGTTCAGTACACAGCTCATCTGAGCTCAGAT	4296
Qy	1068	GlnMetThrGluGlyProAsnThrAlaProProAsnPheSerHisThrGlyProThrPhe	1087	Db	4297	GCCTCTGCTCAGCAGAGACCTCATGATGATAGAGAGACTGTACACACCGAGATCCA	4356
Db	3217	CAATATGAGAGAGACCCAAATACAGCACTTCAAACTTTAGTCTATACGAGCAACATTT	3276	Qy	1448	ThrThrAlaValAlaProSerLysLeuLeuThrSerThrThrSerThrCysLeuProPro	1467
Qy	1088	ProValValProProPheLeuSerSerIleAlaGlyValProThrThrAlaAlaThr	1107	Db	4357	ACTACTCTGTGACCAAGCAAGCACTCTGATCTTACCAACAGTACTGTGTTACCA	4416
Db	3277	CCAGTAGTACTCTCTTTCTTAAAGTACATGCTGTGAGTCCCAACCAAGCAGCAGCCACA	3336	Qy	1468	ThrAsnLeuProLeuGlyThrValAlaLeuProValThrProValValThrProGlyGln	1487

Db 4417 ACCAATTACCACTGAGACAGTCTTCCGCACTTACCACTGAGTCTGAGGACA 4476
 QY 1488 ValSerThrProValSerThrThrThrsSerGlyValLysProGlyThrAlaProSerLys 1507
 Db 4477 GTTCTACCCAGTACGACTACACTACAGAGAGAACTGGAACTGCTCCCTCCAG 4536
 QY 1508 ProProLeuThrLysAlaProValLeuProValGlyThrGluLeuProAlaGlyThrLeu 1527
 Db 4537 CCACCTTACCAAGGCTCCGGTGTGCGACAGTGGTACTGAACTTCCAGCAGTACTTA 4596
 QY 1528 ProSerGluGlnLeuProProPheProGlyProSerLeuThrGlnSerGlnProLeu 1547
 Db 4597 CCGAGCGAGCAGCTCCACCTTTCCAGAGACCTTCTCAACCCAGCTCCAGAGACCTCTA 4656
 QY 1548 GluAspLeuAspAlaGlnLeuArgArgThrLeuSerProGlu**IleThrValThrsSer 1567
 Db 4657 GAGGATCTTGATGCTCAATTGAGAGAACTTACTGACAGATVATCACTGACTTCT 4716
 QY 1568 AlaValGlyProValSerMetAlaAlaProThrAlaIleThrGluAlaGlyThrGlnPro 1587
 Db 4717 GCGGTGCTCTGTCTCCATGCGGCTCCACAGCAATCCAGAGAGAGAGACAGAGCT 4776
 QY 1588 GlnLysGlyValSerGlnValLysGlnGlyProValIleuAlaThrSerSerGlyValaGly 1607
 Db 4777 CAGAAAGGTGTTTCTCAAGTCAAAAGAGCCCTGTCTTACAGCACTAGTTCAAGAGCTGT 4836
 QY 1608 ValPheLysMetGlyArgPheGlnValSerValAlaAlaAspGlyValGlnLysGlnGly 1627
 Db 4837 GTTTTAAAGATGGGAGCATTCAGATTCCTGTTCGACAGACAGCGGCCAGAAAGAGGT 4896
 QY 1628 LysAsnLysSerSerGluAspAlaLysSerValHisPheGluSerSerThrSerGluSerSer 1647
 Db 4897 AAAAATTAAGTCAAGAGATGCAAGATCTGTTCAATTGAAATCCAGCACTTCAGAGTCTCA 4956
 QY 1648 ValLeuSerSerSerSerProGluSerThrLeuValLysProGluProAsnGlyIleThr 1667
 Db 4957 GTGCTATCAGTAGTAGTACAGAGAGTACCTTGTGAAACCAAGAGCCGATGGCATAC 5016
 QY 1668 IleProGlyLysSerSerAspValProGluSerAlaHisLysThrThrAlaSerGluAla 1687
 Db 5017 ATCCCTGATCTCTTCAGATGTGCGACAGAGTGCACAAATCACTGCTCCAGAGCA 5076
 QY 1688 LysSerAspThrGlnGlnProThrLysValGlyArgPheGlnValThrThrAlaAsn 1707
 Db 5077 AAGTAGACACTGGGAGCTTACAGAGTGGACGTTTCAAGTACCACTACAGCAAC 5136
 QY 1708 LysValGlyArgPheSerValSerLysThrGluAspLysIleThrAspThrLysLysGlu 1727
 Db 5137 AAAGTGGGTCGTTCTCTGTATCAAAACTGAGGACAAGATCATGACCAAAAGAA 5196
 QY 1728 GlyProValAlaSerProProPheMetAspLeuGlnAlaValLeuProAlaValIle 1747
 Db 5197 GACCCAGGAGCATCTCTCTCTTATGATTGAAACAAGCTGTTCTCTGCTGTGATA 5256
 QY 1748 ProLysLysGlnLysProGluLeuSerGluProSerHisLeuAsnGlyProSerSerAsp 1767
 Db 5257 CCAAGAAAGAGAGAGCTGAACTGTCAAGCTTCACTTAAATGGGCCGCTCTTGAC 5316
 QY 1768 ProGluAlaAlaPheLeuSerArgAspValAspAspGlySerGlySerProHisSerPro 1787
 Db 5317 CCGGAGGCCGCTTTTAAAGTAGGGAGTGTGATGATGTTCCGGTAGTCAACACCTGCC 5376
 QY 1788 HisGlnLeuSerSerLysSerLeuProSerGlnAsnLeuSerGlnSerLeuSerAsnSer 1807
 Db 5377 CATCAGCTGAGCTCAAAAGAGCTTCTTACGCAAGATCTTAAGTCAAAAGCTTGTATTTCA 5436
 QY 1808 PheAsnSerSerLysMetSerSerAspAsnGluSerAspIleGluAspGluAspLeuLys 1827
 Db 5437 TTAAACTCTCTTACATGATGTCGACATGAGTCAAGATATGAAAGATGAAGACTTAAAG 5496
 QY 1828 LeuGluLeuArgArgLeuArgAspLysHisLeuLysGluIleGlnAspLeuGlnSerArg 1847

Db 5497 TTAGAGCTGGGAGCACTACAGATTAACATCTCAAGAGATTCAAGACTGAGAGTCCG 5556
 QY 1848 GlnLysHisGluIleGluSerLeuThrLysLeuGlyLysValProProAlaValIle 1867
 Db 5557 CAGAGAGCATGAAATTTGATCTTTGTATACCAACTGGGCAAGGGTCCCTGCTGTATT 5616
 QY 1868 IleProProAlaAlaPheLeuSerGlyArgArgArgProThrLysSerLysGlySer 1887
 Db 5617 ATTCCTCCAGCTGTCTCCCTTTCAGGAGAGAGAGAGAGACCACTTAAAGCAAGAGCAGC 5676
 QY 1888 LysSerSerArgSerSerSerLeuGlyValLysSerProGlnLeuSerGlyValLeuSer 1907
 Db 5677 AATCTAGTCGAAGACAGTCTCTTGGGGAATAAAGCCCTTTCAGTAACTGCT 5736
 QY 1908 GlyGlnSerAlaAlaSerValLeuHisProGlnGlnThrLysHisProProGlyValAsnIle 1927
 Db 5737 GGTCAAGTGAAGCTTCACTTGTACCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5796
 QY 1928 ProGluSerGlyGlnAsnGlnLeuGlnProLeuLysProSerProSerSerAspAsn 1947
 Db 5797 CCAGAGTCCGGGAGAGATCAAGCTGTACAGCCCTTAAAGCATCTCCCTCCAGTGAACAC 5856
 QY 1948 LeuTyrSerAlaPheThrSerAspGlyValAlaIleSerValProSerLeuSerAlaProGly 1967
 Db 5857 CTCTATTACGCTTCAACAGATGATGTGTCATTTCAAGTAAAGCTTTCGCTCCAGGT 5916
 QY 1968 GlnGlyThrSerSerThrAsnThrValGlyValAlaThrValAsnSerGlnAlaGlnAla 1987
 Db 5917 CAAGAGACACAGACACAAACAGTGTGGGCAACAGGAAACAGCAAGCCGACCAACT 5976
 QY 1988 GlnProProAlaMetThrSerSerArgLysGlyThrPheThrAspAspLeuHisLysLeu 2007
 Db 5977 CAGCTCTCGCATGACGTCCAGCAGAGAGAGGACATTCAAGATGACTTCACAAAGTTG 6036
 QY 2008 ValAspAsnThrPalaArgAspAlaMetAsnLeuSerGlyArgGlySerLysGlyHis 2027
 Db 6037 GTAGCAATTTGGGCGCGAGATGCCATGATCTCTCAGCAGAGAGAGAGAGAGAGAGAG 6096
 QY 2028 MetAsnTyrGlnGlyProGlyMetAlaArgLysPheSerAlaProGlyGlnLeuCysIle 2047
 Db 6097 ATGAATTATGAGGGCCCTGGAATGCAAGAGAGATCTTGGACTTGGGCAACTGTGATC 6156
 QY 2048 SerMetThrSerAsnLeuGlyGlySerAlaProIleSerAlaAlaSerAlaThrSerLeu 2067
 Db 6157 TCCATGACCTGGAACCTGGGTGGTCTGCTCCCACTCTGAGCATCATCTCTCTA 6216
 QY 2068 GlyHisPheThrLysSerMetCysProProGlnGlnIleGlyPheProAlaThrProPhe 2087
 Db 6217 GGTCACTTCAACAAGTCTATGTGCCCCCAGCAGAGTATGGCTTTCAGCTAACCCCATTT 6276
 QY 2088 GlyAlaGlnTyrSerGlyThrGlyValProAlaProGlnProLeuGlyGlnPheGlnPro 2107
 Db 6277 GAGCTCAATGATGAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 6336
 QY 2108 ValGlyThrAlaSerLeuGlnAsnPheAsnIleSerAsnLeuGlnLysSerLysSerAsn 2127
 Db 6337 GTGGGAATGCGCTCTCTTCAAGATTTCAACATCAAGCAATTTGCAAGAAATCCATGACAA 6396
 QY 2128 ProProGlySerAsnLeuArgThrThr 2136
 Db 6397 CCCCCAGGCTCCAACTTCGAGCACT 6423
 Db
 RESULT 15
 AR302612
 LOCUS AR302612 6885 bp DNA linear PAT 12-JUN-2003
 DEFINITION Sequence 49 from patent US 6541252.
 ACCESSION AR302612
 VERSION AR302612.1 GI:31690910
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 6885)

AUTHORS Walke,D.W., Hilpun,E., Donoho,G. and Turner,C.A. Jr.
 TITLE Human kinases and polynucleotides encoding the same
 JOURNAL Patent: US 5541252-A 49 01-APR-2003;
 FEATURES Location/Qualifiers
 source 1..6885
 /organism="unknown"
 /mol_type="genomic DNA"

ORIGIN

Alignment Scores:

Pred. No.:	3,41e-243	Length:	6885
Score:	10211.00	Matches:	2048
Percent Similarity:	88.20%	Conservative:	0
Best Local Similarity:	88.20%	Mismatches:	0
Query Match:	94.44%	Indels:	274
DB:	6	Gaps:	2

US-10-010-720-14 (1-2136) x AR302612 (1-6885)

QY 61 MetAspLysAspSerArgGlyAlaAlaAlaThrThrThrThrGluHisArgPhe 80
 Db 1 ATGGAACAGAGACGCGGTGGGGCGCGCCGACACCACTACCACTAGAGACCGCTTCTT 60
 QY 81 ArgArgSerValIleCysAspSerAsnAlaThrAlaLeuGluMetProGlyLeuProLeu 100
 Db 61 CCGCGAGCGTCACTGCGACTCCATGCGACTGAGAGCTTCCGAGCTTCTCTT 120
 QY 101 SerLeuProGluProSerIleProAlaAlaValProGluSerAlaProProGluProHis 120
 Db 121 TCCCTGCCCCAGCCGACATCCCGCGGCTGTCGCCGACAGAGCTCCACGAGCCCCAC 180
 QY 121 ArgGluGluThrValThrAlaThrAlaThrSerGluValAlaGluGluProProAlaAla 140
 Db 181 CCGGAAGAGACCGGTGACCGCCACCGCCACTTCCAGTAGCCAGACGCTCCAGCGCT 240
 QY 141 AlaAlaProGlyGluGluAlaValAlaGlyProAlaProSerThrValProSerSerThr 160
 Db 241 GCGGCGCCCTGGGGAACAGGCCCTCGCGGCTTCCGCTCCACTGTCCTCCAGCACTACC 300
 QY 161 SerLysAspArgProValSerGluProSerLeuValGlySerLysGluGluProProPro 180
 Db 301 AGCAAAAGACCGCCGAGTCTCCAGCCTTGTGGGAGAGAAAGAGAGCCGCCCGC 360
 QY 181 AlaArgSerGlySerGlyGlyGlySerAlaLysGluProGluGluGluArgSerGluGlu 200
 Db 361 GCGAAGATGGGAGCGCGCGCGCCAGCCAGGAGCCAGAGGAGAGAGAGAGAGAGAGAG 420
 QY 201 GluAspSerIleGluGluLeuGluThrLysAlaValGlyMetSerAspArgLysPhe 220
 Db 421 CAGATGATATCGAAGAGCTGAGACCAAGGCCGTGGAAATGCTTAAGATGGCGCTT 480
 QY 221 LeuLysPheAspIleGluIleGlyArgGlySerPheLysThrValTyrLysGlyLeuAsp 240
 Db 481 CTCAAGTTTGACATCGAAATCGGACAGAGGCTCTTTAAAGCGCTCAAAAGCTGAGC 540
 QY 241 ThrGluThrThrValGluValAlaAlaTyrCysGluLeuGluAspArgLysLeuLysSer 260
 Db 541 ACTAAACCAACCGTGAAGTGGCTGTGATGAGTCAAGATCGAAATTAACAAGAGCT 600
 QY 261 GluArgGluArgPheLysGluGluAlaGluMetLeuLysGlyLeuGluHisProAsnIle 280
 Db 601 GAGAGGACAGAGATTAAAGAAAGAGCTGAATGTTAAAGGTCTTCAAGATCCCAATATT 660
 QY 281 ValArgPheLysAspSerTyrGluSerThrValLysGlyLysCysIleValLeuVal 300
 Db 661 GTTAGATTATATATCTCTGGAAATCCACAGTAAAGAAAGAGATGCTTTTGGT 720
 QY 301 ThrGluLeuMetThrSerGlyThrLysThrThrTyrLeuLysArgPheLysValMetLys 320
 Db 721 ACTGAACCTTGACGCTGGAACAATTAAACGATATCGAAAGGTTTAAAGATGAGAG 780
 QY 321 IleLysValLeuArgSerTyrCysArgGluIleLeuLysGlyLeuGluPheLeuHisThr 340

Db 781 ATCAAACTCTAAGAGAGCTGGTGGCCGACAGATCCTTAAGCTTCTTCACTACT 840
 QY 341 ArgThrProProIleIleHisArgAspLeuLysCysAspAsnIlePheIleThrGlyPro 360
 Db 841 CGAAGCTCACTATCATTTACCCCGAGATCTTAATATGTAACAACATCTTATCACCGGCCCT 900
 QY 361 ThrGlySerValLysIleGlyAspLeuGlyLeuAlaThrLeuLysAlaGluAspPheAla 380
 Db 901 ACTGGCTCACTCAAGATTTGAGAGCTGGCTTGACCAACCTCGAAGCGGCTTCTTGGCC 960
 QY 381 LysSerValIleGlyThrProGluPheMetAlaProGluMetTyrGluGluLysTyrAsp 400
 Db 961 AAGAGTGATATGATGCCCAAGATTCAAGCTTCAAGCTTCAAGATGTATGAGGAATATGAT 1020
 QY 401 GluSerValAspValTyrAlaPheGlyMetCysMetLeuGluMetAlaThrSerGluTyr 420
 Db 1021 GAATCGGTGACGTTATGCTTTGGAGTGTGATGATGCTTGAAGTGCATCTGAATAT 1080
 QY 421 ProTyrSerGluCysGluAsnAlaAlaGluIleTyrArgArgValThrSerGlyValLys 440
 Db 1081 CCTTACTCGAGTGCCTAAATCTGACAGATCTTACCGCTGACACAGTGGGGTGAGAG 1140
 QY 441 ProAlaSerPheAspLysValAlaIleProGluValLysGluIleIleGluGlyCysIle 460
 Db 1141 CCAGCCAGTTTGAACAAAGTACCATTCCTGAGAGTGAAGAAATTTTGAAGAGATGCATA 1200
 QY 461 ArgGluAsnLysAspGluArgTyrSerIleLysAspLeuLysAsnIleAlaPhePheGlu 480
 Db 1201 CGACAAACAAAGAGAAAGATATTCATCAAAAGACCTTTTGAACATGCTTCTTCCA 1260
 QY 481 GluGluThrGlyValArgValGluLeuAlaGluGluAspArgLysIleAlaIle 500
 Db 1261 GAGGAACAGAGATGCGGTGAAATTAACAGAAAGATGATGGGAAATAATATACCATTA 1320
 QY 501 LysLeuTyrPheLysArgIleGluAspIleLysLysLysLysGlyLysTyrLysAspAsnGlu 520
 Db 1321 AAATTATGCTACGATTAATTAAGAAATTTAAAGGAAATTAACAAAGATATAGAA 1380
 QY 521 AlaIleGluPheSerPheAspLeuGluArgAspValProGluAspValAlaGluGluMet 540
 Db 1381 GCTATTGATTTTCTTTTGAATTTAGAGAGATGCCAGAAAGATGTTGCCACAGAAATG 1440
 QY 541 ValGluSerGlyTyrValCysGluGlyAspHisLysThrMetAlaLysAlaIleLysAsp 560
 Db 1441 GTAGAGTGTGGATGATGCTGAAAGGTGATCAACAGACCATGGCTTAAAGTATCAAGAC 1500
 QY 561 ArgValSerLeuIleLysArgLysArgGluGluArgGluLeuValArgGluGluGluGlu 580
 Db 1501 AGAGTATCATTAATTAAGAGAAACGAGACGCGAGTTGGTACGAGAGAGCAAGAA 1560
 QY 581 LysLysLysGluGluGluSerSerLeuLysGluGluValGluGluSerSerAlaSerGlu 600
 Db 1561 AAAAAAAGAGAGAGAGACAGTCAACAGAGGTAGAACATTCAGTCTTCCAG 1620
 QY 601 ThrGlyLysLeuGluLeuProSerAlaSerThrGlyLysProThrAlaSerThrThrSer 620
 Db 1621 ACGAGATCAAGAGCTCCCTTCTGAGACCGGCACTACCTGCTTCTTACCACTTCA 1680
 QY 621 AlaSerValSerThrGluValGluProGluGluProGluAlaAspGluHisGluGluLeu 640
 Db 1681 GCTTCAAGTTCTACACAGTGAACCTGAAGAACCTGAGGCAATCAACATCAACACTA 1740
 QY 641 GluTyrGluGluProSerIleSerValLeuSerAspGlyThrValAspSerGlyGluGly 660
 Db 1741 CAGTACAGCAACCCAGATATCTGTATCTGATGAGCGGTGACAGTGTGACAGGA 1800
 QY 661 SerSerValPheThrGluSerArgValSerSerGluGluThrValSerTyrGlySerGlu 680
 Db 1801 TCCCTGCTTCAACAATCTCAGTGAAGAGCAACAAGCAATTTCAATATGTTCCCA 1860
 QY 681 HisGluGluAlaHisSerThrGlyThrValProGluHisIleProSerThrValGluAla 700
 Db 1861 CATGAACAGGACATCTACAGGACAGTCCAGGCAATATACCTTCTACTGTCCAGCA 1920

QY	701	GlnSerGlnProHisGlyValTyrProProSerSerValAlaGlnGlnSerGlnGly	720	QY	815	ProThrThrLeuAlaSerSerValAspSerAlaHisSerAspValAlaSerGlyMetSer	834
DB	1921	CAGCTCAGGCCCATGGGGATATATCCACCTCAAGTGTG-----	1959	DB	2917	CCGACCACCTTGGCTTCTCTGTAGACATGACATTCAGATGTGCTTCAAGTATAGT	2976
QY	721	GlnProSerSerSerSerLeuThrGlyValSerSerSerGlnProIleGlnHisProGln	740	QY	835	AspGlyAsnGlnLeuValProSerSerSerGlyArgHisGlnGlyArgThrThrLeuArg	854
DB	1959	-----	1959	DB	2977	GATGGCAATGAGAACGCCATCTTCAGTGGAGGAGGTATAGGAACTACCAAAACGG	3036
QY	741	GlnGlnGlnGlnGlyIleGlnGlnThrAlaProProGlnGlnThrValGlnTyrSerLeuSer	760	QY	855	HisTyrArgIleSerSerValArgSerArgSerArgHisGlnIleThrSerArgProIleLeu	874
DB	1960	---CAGCAGGGAATACAGCAGACAGCCCTCTCCACAGCAGAGTCACTTCA	2016	DB	3037	CATTACCGAAATCTGTAGAGAGCGCTCTCGACATGAAAAAATTCACGCCCAAAATTA	3096
QY	761	GlnThrSerThrSerSerSerGlyAlaThrThrAlaGlnProValSerGlnProGlnAlaPro	780	QY	875	ArgIleLeuAsnValSerAsnIleGlyAspArgValAlaGlnCysGlnIleGlnThrHis	894
DB	2017	CAGACATCAACCTCCAGTGAAGGCCACTACTGCACAGCCAGAGTCAAGCTTCA	2076	DB	3097	AGAAATTTGATGATTTCAATTAAGGAGACCGAGTATGATGATGATTAAGAGCTCAT	3156
QY	781	GlnValLeuProGlnIleValSerAlaGlyIleGln-----	791	QY	895	AsnArgIleSerValIleThrPheIlePheAspLeuAspGlyAspAsnProGlnGlnIleAla	914
DB	2077	CAGCTTGTGCTCAAGTATACAGCTGGAAACAGCTTCCAGTTCCAGCCAGTACCAACT	2136	DB	3157	AATGGAATAATGTTACATTCAAATTTGACTTAATGTGTGCACAAACCCGAGGAGATNGCA	3216
QY	791	-----	791	QY	915	ThrIleMetValAsnAsnAspPheIleValAlaIleGluArgGlnSerPheValAspGln	934
DB	2137	ATCCAGGCGCAACCTCAGATCCAGTGTGCGACACAAACCTCGTGTGTTCCAGTCCACTCT	2196	DB	3217	ACAAATTAATGGAACAAATGACTTTATTTAGCAATAGAGAGAGAGTCTGTGTGATCA	3276
QY	791	-----	791	QY	935	ValArgGlnIleIleGlnIleValAlaAspGlnMetLeuSerGlnAspValSerValGlnPro	954
DB	2197	GGTGCTACTTTCTTTCAGTGGGAGACGCCCTCCTACTCCTTGTCTCCCTCAGTACCT	2256	DB	3277	GTGGCAGAAATTAATGAAAAAGCTGATGATAATGCTCACTGAGATGATGATGATGACCA	3336
QY	791	-----	791	QY	955	GlnGlyAspGlnGlyLeuGlnSerIleGlnIleGlyAspAspIleGlyPheSerGlySer	974
DB	2257	GTTCTCTCAGATTCCTCATATCAACTCTCTCATGTGTCTAGCGCTCAGACAGGTTTCTCATCC	2316	DB	3337	GAGGTGATCAGGGATGAGAGTGTACAGAGAAAGATGACTATGCTTTTTCAGGTTCT	3396
QY	791	-----	791	QY	975	GlnIleSerLeuGlnGlyGlnPheIleGlnProIleProAlaSerSerMetProGlnIle	994
DB	2317	CTTCCCATCACAATGGAGCTGGCATTTACTCAGCTCTGCTCAGCTTGGCTTCACTGTCT	2376	DB	3397	CAGAAATTTGAAAGAGAGATTCACAAACCAATTCCTGGCTCTTCATGCCACAGCAAAATA	3456
QY	791	-----	791	DB	995	GlyIleProThrSerSerSerLeuThrGlnValAlaHisSerAlaGlyArgPheIleVal	1014
DB	2377	ACAAAGCTGCGATCCCGGGGATATCACTGTGTCTTCTAGTCACTTCAACCTTCTTG	2436	DB	3457	GGCATTCCTACAGATTTCTTAACTCAAGTGTTCATTTGCGCGGAAGGGGTTTATAGT	3516
QY	791	-----	791	QY	1015	SerProValProGlnSerArgLeuArgIleSerIleValPheProSerGlnIleThrAsp	1034
DB	2437	TCCATGGGAATACAGCTAACTTGGACAAGCTGTGAGGTTCCACTTCTCTGGAGAT	2556	DB	3517	ACAGTTCCTGCTCTTACAGCTCAGAGCCCTGGAATAACTGTCTCATCTGCATCATCC	3636
QY	791	-----	791	QY	1055	LeuSerLeuGlnGlnAlaPheSerGlnLeuAlaArgAlaGlnMetThrGlnGlyProAsn	1074
DB	2557	GTTCTGTACCAAGGCTTCCCACTCGACTGCACACACAGTACCCAGAGATTCAAATATTT	2616	DB	3637	CTTAGCTTACCAACAGGCTTTTCTGAACTTAACTGACGCTCAATGACAGAGAGACCAAT	3696
QY	791	-----	791	QY	1075	ThrAlaProProAsnPheSerHisThrGlyProThrPheProValValProProPheLeu	1094
DB	2617	GCTCCCTCTTCCAGCTGCTTCTGTGTGATCATTTCTACAGTCTCTAMCCCTCCCATG	2676	DB	3697	ACAGCACCTCCAACTTTAGTCTATACAGACCAACATTTCCAGTACTACTCTTCTTA	3756
QY	791	-----	791	QY	1095	SerSerIleAlaGlyValProThrThrAlaAlaAlaThrAlaProValProAlaThrSer	1114
DB	2677	CCGACAGAACTACTGGCTACACCTGGGTACTTCCACAGTGTGACGCTTATATGGAA	2736	DB	3757	AGTAGATGCTGGAGATGCCCAACACACAGCAGCAGCAGCAGCAGCTCTGCAACAGC	3816
QY	791	-----	791	QY	1115	SerProProAsnAspIleSerThrSerValIleGlnSerGlnValThrValProThrGlu	1134
DB	2737	TCAAACTTTTATTTCTATGGGTGTGTAGAGAGACAGGTTCAGTGTCCCAAGCA	2796	DB	3817	AGCCCTCTTAATACATTTTACATACAGACCAACATTTCCAGTACTACTCTTCTTA	3876
QY	792	-----	792	QY	1135	GlnGlyIleAlaGlyValAlaThrSerThrGlyValValThrSerGlyIleProIle	1154
DB	2797	GGGAGTTTACACAGCCCACTACATCTCTCCACAGCAGAGTTTGGAGAGTACTCAG	2856	DB	3877	GAGGGAGTTGCTGAGTGTCCACACAGCAGAGTGTGTACTTCAAGTGTCTCCCATATA	3936
QY	795	GlyValSerGlnValAlaProAlaGluProValAlaValAlaGlnProGlnAlaThrGln	814	QY	1155	ProProValSerGlnSerProValIleSerSerValIleSerSerIleThrIleProAla	1174
DB	2857	GGAGTCTCTCAGGTGTGCTCTCGCAGAGCAGTGTGAGTACAGCAGCCCAAGCTACCCAG	2916	DB	3937	CCACTGTGTCTGATATACAGATACCTTCCAGGCTAGTTCAGAGTATCACAAATACCTGCA	3996
QY				QY	1175	ValValSerIleSerThrThrSerProSerLeuGlnValProThrSerThrSerGlnIle	1194

Db 3997 GTTGTCTAATATCTACTACATCCCGCTCACTTCAAGTCCCAATCCACATCTGAGATC 4056
Qy 1195 ValValSerSerThrAlaLeuTyrProSerValThrValSerAlaThrSerAla 1214
Db 4057 GTTCTTCTAGTACAGCACTGATCTTCAAGTACAGTTTCAAGCACTTCAAGCTTCA 4116
Qy 1215 GlyGlySerThrAlaThrProGlyProLysProProAlaValValSerGlnAlaAla 1234
Db 4117 GGGGGGAGTACTGCTACCCAGGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCA 4176
Qy 1235 GlySerThrThrValGlyAlaThrLeuThrSerValSerThrThrThrSerPhePro 1254
Db 4177 GAGAGCACTACTGTGGAGGACCATTAATCACTTCAACCACTTCACTTCACTTCA 4236
Qy 1255 ThrAlaSerGlnLeuSerIleGlnLeuSerSerSerSerThrProThrIleuAlaGlu 1274
Db 4237 ACAGCTTCAAGCTGCTCACTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAG 4296
Qy 1275 ThrValValValSerAlaHisSerLeuAlaSerThrSerThrThrThrThrThr 1294
Db 4297 ACCGGTAGTATAGGCACTCACTCAATGATAGACATCTCAATGACAGTACAGTATG 4356
Qy 1295 AlaPheSerLeuSerAlaProSerSerSerSerSerProGlyAlaGlyValSerSer 1314
Db 4357 GCTTCTTCCCTCTCTGACCATCTTCTCTTCTCTCTCTCTCTCTCTCTCTCTCTCT 4416
Qy 1315 IleSerGlnProGlyGlyLeuHisProLeuValIleProSerValIleAlaSerThr 1334
Db 4417 ATTCTTACGCTGCTGGGCTGACATCTTGTGATTCATCAATGATAGTATGCTTCA 4476
Qy 1335 IleLeuProGlnAlaAlaGlyProThrSerThrProLeuLeuProGlnAlaProSer 1354
Db 4477 ATTCTTCCCAAGCAGACAGCACTTCTTCACTTATTAACCCCAAGTACCTATGATC 4536
Qy 1355 ProProLeuValGlnProValAlaSerValProAlaValGlnThrLeuIleHisSer 1374
Db 4537 CCACCTTGGTACAGCTGCTTCCCAATGCTGCTGTAACAGCAACATTAATTCATAGT 4596
Qy 1375 GlnProGlnProAlaLeuLeuProGlnGlnProHisSerHisSerProGlnValAla 1394
Db 4597 CAGCTTCAACCAAGCTTCTTCCCAAGCAGCCTTCACTTCTTCTTCTTCTTCTTCT 4656
Qy 1395 AspThrGlnProLysAlaProGlyIleLeuAspIleLeuGlnGlnGlnGlnGlnGln 1414
Db 4657 GATACACACCAACCAAGCTTCTTCAATGATGATGATGATGATGATGATGATGATGAT 1414
Qy 1415 SerLeuPheSerGlnHisSerSerSerSerGlyAlaGlnHisAlaSerValSerLeu 1434
Db 4717 TCTCTGTTCAGTGAACACAGCTCATCTGAGCTTCAAGCATGCTGCTCACTGAGAAC 4776
Qy 1435 SerLeuValIleGlnSerThrValThrProGlyIleProThrThrAlaValAlaPro 1454
Db 4777 TCACTGATCATAGAGCACTGTCACACGAGCATCCCACTTCTCTGTCAGTGAAGAC 4836
Qy 1455 LysLeuLeuThrSerThrThrSerThrCysLeuProProThrAsnLeuProLeuGln 1474
Db 4837 AAACCTCTGATCTTCAACCAAGTACTTCTTCAACCAACCAATTTACCACTAGAGAAC 4896
Qy 1475 ValAlaLeuProValThrProValValThrProGlyGlnValSerThrProValSer 1494
Db 4897 GTTGTCTTGGCCAGTACACAGCTGCTCACTGAGCAAGTTTCAACCCAGTCAAGCT 4956
Qy 1495 ThrThrSerGlyValLysProGlyThrAlaProSerLysProProLeuThrLysAlaPro 1514
Db 4957 ACTACATAGAGAGGAACTGGAAGCTGCTCTCTCAAGCCCACTTCAACCAAGCTCGC 5016
Qy 1515 ValLeuProValGlyThrGlnLeuProAlaGlyThrLeuProSerGlnGlnLeuProPro 1534
Db 5017 GTGCTGCAAGTGGTACTGAACTTCCAGAGTACTTCAACCAAGCAGAGCTGCAAGCT 5076
Qy 1535 PheProGlyProSerLeuThrGlnSerGlnGlnProLeuAlaSerValAlaGlnLeu 1554
Db 5076 GTTGTCTAATATCTACTACATCCCGCTCACTTCAAGTCCCAATCCACATCTGAGATC 4056
Qy 1555 ArgArgThrLeuSerProGlu***IleThrValThrSerAlaValGlyProValSer 5137
Db 5137 AGAAGAACACTTGTAGTACAGATATATCAGTACTTCTGCGGTGATCTCTGATCAGT 5175
Qy 1575 AlaAlaProThrAlaIleThrGlnAlaGlyThrGlnProGlnLysGlyValSerGln 5197
Db 5197 GCGGCTTCCAAAGCAATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5257
Qy 1595 LysGlnGlyProValLeuAlaThrSerSerGlyAlaGlyValPheLysThrGlyArgPhe 5257
Db 5257 AAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1615
Qy 1615 GlnValSerValAlaAlaAspGlyValGlnLysGlnGlyLysAsnLysSerGlnAla 1634
Db 5317 CAGGTTCTGTTGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1635
Qy 1635 LysSerValHisPheGlnSerSerThrSerGlnSerValLysSerSerSerPro 1654
Db 5377 AAGTCTGTTCAATTTGAATCCAGCACTCAAGATCTTCAAGTATCAAGTATGATCA 1674
Qy 1655 GluSerThrLeuValLysProGlnProGlnGlyIleThrIleProGlyIleSerSer 1674
Db 5437 GAGAGTACCTTGTGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1694
Qy 1675 ValProGlnSerAlaHisLysThrThrAlaSerGlnAlaLysSerAspThrGlnPro 1694
Db 5497 GTGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1714
Qy 1695 ThrLysValGlyArgPheGlnValThrThrThrAlaAsnLysValGlyArgPheSer 1714
Db 5557 ACCAAGGTGAGAGGTTTCAAGTGAACATCAAGCAACCAAGAGAGAGAGAGAGAG 1734
Qy 1715 SerLysThrGlnAspLysIleThrAspThrLysGlnGlyProValAlaSerProPro 1734
Db 5617 TCAAAATCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1754
Qy 1735 PheMetAspLeuGlnAlaValLeuProAlaValIleProLysGlnLysProGln 1754
Db 5677 TTTATGATTTGGAACAGAGCTGTTCTTCTCTCTGATGATGATGATGATGATGAT 1774
Qy 1755 LeuSerGlnProSerHisLeuAsnGlyProSerSerSerProGlnAlaAlaPheLeu 1774
Db 5737 CTGTCAAGAGCTTCAACATTAATGAGCCCTTCTTCAAGCCGAGGCGCTTTTAACT 1794
Qy 1775 ArgAspValAspAspGlySerGlySerProHisSerProHisGlnLeuSerLysSer 1794
Db 5797 AGGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1814
Qy 1795 LeuProSerGlnAsnLeuSerGlnSerLeuSerSerPheAsnSerSerThrMetSer 1814
Db 5857 CTTCTTCAAGCAATTAATGATCAAGCTTCAATTAATTAATTAATTAATTAATTAAT 1834
Qy 1815 SerAspAspGlnSerAspIleGlnAspGlnSerArgGlnLysHisGlnIleGlnSer 1834
Db 5917 AGCGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1854
Qy 1835 AspLysHisLeuLysGlnIleGlnAspLeuGlnSerArgGlnLysHisGlnIleGlnSer 1854
Db 5977 GATTAACATCTCAAGAGATTCAGAGCTGCAAGAGTCCCAAGAGATGAAATTAATCT 1874
Qy 1855 LeuTyrThrLysLeuGlyLysValProProAlaValIleIleProProAlaAlaPro 1874
Db 6037 TTGTATACCAAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1894
Qy 1875 SerGlyArgArgArgArgProThrLysSerLysGlySerLysSerSerArgSerSer 1894
Db 6097 TCAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1914
Qy 1895 LeuGlyAsnLysSerProGlnLeuSerGlyAsnLeuSerGlyGlnSerAlaAlaSerVal 1914
Db 6157 TTGGGAAATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6216

QY 1915 LeuHisProGlnGlnThrLeuHisProProGlyAsnIleProGlnSerGlyGlnAsnGln 1934
 |||||
 DB 6217 TTGGACCCCGAGAGACCCCTCCCTGCGACATCCGAGAGTCCGGGAGAGATCAG 6276
 |||||
 QY 1935 LeuLeuGlnProLeuLysProSerProSerSerAspAsnLeuTyrSerAlaPheThrSer 1954
 |||||
 DB 6277 CTGTTACAGCCCTTAAGCCATCTCCCTCAGTGACACCTCTATTCAGCCTTCAACAGT 6336
 |||||
 QY 1955 AspGlyAlaIleSerValProSerLeuSerAlaProGlyGlnGlyThrSerSerThrAsn 1974
 |||||
 DB 6337 GATGTGCGCATTCATGATCAGACCTTTCTGCTCCAGGTCAAGAACCGAGACACAAAC 6396
 |||||
 QY 1975 ThrValGlyAlaThrValAsnSerGlnAlaAlaGlnIleGlnProProAlaMetThrSer 1994
 |||||
 DB 6397 ACTGTTGGGCGACAGTGAACAGCAAGCCGCGCCAGCTCAGCTCCTGCGCATGAGCTCC 6456
 |||||
 QY 1995 SerArgLysGlyThrPheThrAspAspLeuHisLysLeuValAspAsnTrpAlaArgAsp 2014
 |||||
 DB 6457 AGCAGGAAGGCGCATTCACAGATGACTTGCAACAAGTTGTAGACAATTGGGCCCGAGAT 6516
 |||||
 QY 2015 AlaMetAsnLeuSerGlyArgArgGlySerLysGlyHisMetAsnTyrGlnGlyProGly 2034
 |||||
 DB 6517 GCCATGATCTCTCAGGAGAGAGAGCAAGGCGACATGATGATGAGAGGCGCTTGA 6576
 |||||
 QY 2035 MetAlaArgLysPheSerAlaProGlyGlnLeuCysIleSerMetThrSerAsnLeuGly 2054
 |||||
 DB 6577 ATGGCAAGGAAGTTCTGCGACCTGGGCAACTGTGCATCTCCATGACCTCGAACCTGGGT 6636
 |||||
 QY 2055 GlySerAlaProIleSerAlaIleSerAlaThrSerLeuGlyHisPheThrLysSerMet 2074
 |||||
 DB 6637 GGCTGCGCCCATCTCTGCGATCAGCTACTCTCTAGTCACTTCAACAAAGTCTATG 6696
 |||||
 QY 2075 CysProProGlnGlnTyrGlyPheProAlaThrProPheGlyAlaGlnTrpSerGlyThr 2094
 |||||
 DB 6697 TGCCTCCCGACAGCATGTGGCTTCCAGTACCCCAATTTGGCGCTCAATGAGTGGAGC 6756
 |||||
 QY 2095 GlyGlyProAlaProGlnProLeuGlyGlnPheGlnProValGlyThrAlaSerLeuGln 2114
 |||||
 DB 6757 GGTGGCCCGACAGCACAGCACTTGCGCAGTCCACACTGTGGGAACCTGCTCTTGAG 6816
 |||||
 QY 2115 AsnPheAsnIleSerAsnLeuGlnLysSerIleSerAsnProProGlySerAsnLeuArg 2134
 |||||
 DB 6817 AATTCAACATCAGCAATTGCGAATCATCAGCAACCCCGAGGCTCCCACTGCGG 6876
 |||||
 QY 2135 ThrThr 2136
 |||||
 DB 6877 ACCACT 6882
 |||||

Search completed: September 25, 2004, 08:40:52
 Job time : 12502 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 24, 2004, 01:00:53 ; Search time 64 Seconds
(without alignments)
3210.392 Million cell updates/sec

Title: US-10-010-720-14
Perfect score: 10812
Sequence: 1 MSGAAEKOSSTPGSLFLSP.....NISNLOKISINPPGSNLRRT 2136

Scoring table: BROSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1656	15.3	1851	2 T19964	hypothetical prote
2	765.5	7.1	633	1 S49611	probable serine/th
3	745.5	6.9	516	2 T06716	hypothetical prote
4	702.5	6.5	547	2 T46059	MAP kinase - Arabi
5	702.5	6.5	549	2 T51039	MAP kinase (import
6	678.5	6.3	677	2 T02951	probable mitogen a
7	591.5	5.5	2232	2 T34434	hypothetical prote
8	494	4.6	3570	2 T45025	mucin MUC5B, trach
9	478	4.4	1802	2 S69703	HK1 protein precu
10	475.5	4.4	2090	2 S26058	probable transform
11	474.5	4.4	1952	2 T48814	hypothetical prote
12	472	4.4	2187	2 T30826	nascent polypeptid
13	468	4.3	1192	2 T18611	probable serine/th
14	467.5	4.3	3507	2 T34513	hypothetical prote
15	460	4.3	1367	1 S48478	hypothetical prote
16	457	4.2	1398	2 T13741	hypothetical prote
17	454.5	4.2	1246	2 G89287	glucan 1,4-alpha-g
18	434.5	4.0	2271	2 F90073	protein H39E23.1
19	431.5	4.0	792	2 JC7122	hypothetical prote
20	428.5	4.0	1777	2 E95206	hypothetical prote
21	428	4.0	4776	2 E95206	cell wall surface
22	421.5	3.9	13288	2 T03099	mucin, submaxillar
23	416.5	3.9	1630	2 A53577	ascites staloglyco
24	416	3.8	1080	2 S48944	hypothetical prote
25	416	3.8	5327	2 T13564	microtubule-associ
26	413	3.8	1459	2 T32271	hypothetical prote
27	406	3.7	1829	2 T24583	hypothetical prote
28	403	3.7	2529	2 A56923	transcription fact
29	398	3.7	1233	2 T30989	serine/threonine p

30	396.5	3.7	528	2 I47141	gastric mucin (c10
31	396.5	3.7	1734	2 A54602	microtubule-associ
32	393.5	3.6	2578	2 A56922	transcription fact
33	388.5	3.6	865	2 A47282	calcium-binding pr
34	387.5	3.6	1306	2 S25370	MSB2 protein - yea
35	385.5	3.6	2464	1 QRMSP1	microtubule-associ
36	385	3.6	841	1 I76885	serine/threonine-s
37	383	3.5	2264	2 A56577	microtubule-associ
38	383	3.5	2453	2 S60254	nuclear receptor c
39	382.5	3.5	1609	2 S25345	probable membrane
40	381.5	3.5	1611	2 T38236	hypothetical prote
41	380	3.5	1275	2 T33369	hypothetical prote
42	378	3.5	939	2 S28394	probable serine/th
43	374	3.5	1206	2 T34021	protein kinase SK2
44	373.5	3.5	873	2 A47283	calphostin - fruit
45	373.5	3.5	1032	2 T34433	hypothetical prote

ALIGNMENTS

RESULT 1

T19964 hypothetical protein C46C2.1 - Caenorhabditis elegans

C/Species: Caenorhabditis elegans

C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C/Accession: T19964

R/McMurray, A. submitted to the EMBL Data Library, December 1995

A/Reference number: Z19204

A/Accession: T19964

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-1851 <WIL>

A/Cross-references: EMBL:Z68296; P1DN:CAA92591.1; GSPDB:GN00022; CESP:C46C2.1

A/Experimental source: clone C46C2

C/Genetics:

A/Gen: CESP:C46C2.1

A/Map position: 4

A/Introns: 15/3; 42/1; 65/2; 93/3; 239/2; 340/1; 448/3; 863/3; 1096/3; 1351/3; 1495/3; 1

Query Match 15.3%; Score 1656; DB 2; Length 1851;

Best Local Similarity 27.1%; Pred. No. 1.5e-41;

Matches 605; Conservative 272; Mismatches 646; Indels 710; Gaps 87;

QY	8	KQSTPGSLFLSPAPAPKXGSSDSVG-----EKLGAADAVTGR--50
DB	14	RMGITWCG--RPPAPSSVSTTASTGNGFRRLVNRIRKYVDLHPAQENPTWGSW 70
QY	51	-TEYRRRRRTMDKSGAATTTTTEHR-----FFRSYICDNATALEPGL--PLSL 102
DB	71	LSBEGRSLRAVQDDMKRTPKMTSKKRPDDPTTSSPSVLSNALENSTPSINNVSSI 130
QY	103	PGSPISPAVPSAPPEBHEEYVATA---TSQVAQOP-----138
DB	131	TNSSSPFSLGAA-----TSTASAIIPFTSVATNPHLNHVRIPQAVTGTNG 182
QY	139	-----AAAPGQAVAP-APST-----155
DB	183	SLPPLLSPTSAATPLISGKAGPMSPSTPINVAATVQNAVSSPHSIFDRSRINK 242
QY	156	VPSSTG-----KDRVSGPSLVGSK-----EPPPARSG-----SGG 187
DB	243	IPPNSTLASSSSPSDANNKPIQORHSILSNVRLTQAMVNDGRTLTGDDMDQVSEE 302
QY	168	GSAPKQPERSQOQD-----DIE---ELDYAVGNSDGRFLKFDLELGRGSKFYVYGL 239
DB	303	EPARKQKREBEHGAARIDVEDDFDAQEKPIDSKNGRFLKFPDELELGRGSKFYVYGL 362
QY	240	DTETTEVAVMCELQDRKLTSEKQRFKFAEMLKGLOHNIVAFVDSWST--VKGR-KCI 297
DB	363	DIETGVAVAMCELQSKINKTERQRFREAEMLKDLQHNIVAFVDFYWSADLCGRKXI 422

QY 298 VLVTLEMTSGTLKTYLKEFKVMKIR-VLRSMCRQIIKGLQFLHTRPPIIHDKCNIE 356
 Db 423 VLVTLEMTSGTLKTYLKEFKVMKIR-VLRSMCRQIIKGLQFLHTRPPIIHDKCNIE 482
 QY 357 ITGPTGSVYIGLGLATTLKRAFPKSVIGTPEPMPEMEEEKYDESVDVYAFGMCLMENA 416
 Db 483 ITGPTGSVYIGLGLATTLKRAFPKSVIGTPEPMPEMEEEKYDESVDVYAFGMCLMENA 542
 QY 417 TSEYVYSECCNAQAQIRRVYTSVGVKPAFEDKV-AIPVEKEITEGCIRKQKDRYSIKUL 474
 Db 543 TSEYVYSECCNAQAQIRRVYTSVGVKPAFEDKV-AIPVEKEITEGCIRKQKDRYSIKUL 602
 QY 475 NHAFFQEE--TGVRYLEAED--DGEKIALKMLRIDIKKLG-KYKXNEAIEFSDI 528
 Db 603 VDDEFTPEBDLIGIRVEIKRNDADLNDLVEITOMQLKVDDEKKRQYREKNEGQIFADI 662
 QY 529 ERDVPEDVAQMEVSGVYCEGDHKTMAKIKDRVSL-----IKRREORQVRE 577
 Db 663 ENDSPPDEVQOMIEQOHIPDEDTMITKLIDKIDAFRRDHRDLLETKRKEEERIRE 722
 QY 578 EQE-----KKQOESSLKOQVEQSSASQGTGKOLPSASTGIPTASTTSASTQ 626
 Db 723 EAEIKELRLRAEKERKEKEKLEKRAAAANPNP--TPIPTPATPHS-SAQ 778
 QY 627 VEDEEP--EADQHQLOQYQOQPSISVLSQGTVDNSGQSSVFTESRVSSQOTVSYGQHE 682
 Db 779 QQPIPPPLSTQTAHIOQSAQOQSVV----- 805
 QY 683 QAHSTGVFEGHISTVQAOQOPHGVYPPSSVA-----QGSQOQSSSLT 728
 Db 806 -----TMTANIPAMSPSAQOPVLSPTISAAPVPTMTIHVKEPEIPVQVATTAAPV 859
 QY 729 GVSSQPIQHPOQOQIQTAPPOQTVQVSLQSTSTSEATTAQVPSQAPQVLPQVSA 788
 Db 860 AANNPPSPAPEFTEDIQPTLAQNTVPRTISTDAGLVNTPASTASP-----SP 910
 QY 789 GKOSTQGVQAFAEPAVAQAQOQPTPLASVDSASDASGMSGDNENVPSSGRHE 848
 Db 911 APSATDVASTAPVT-----APPTT-----TDDGAAASTTTEK 948
 QY 849 GNTYKTRHRSVSSSRHEKTRPKILNV-SNKQDRVVEGOLTHNKMTKPFEDLDG 907
 Db 949 EKRKSNRKXV-----MELIGCESNFPALVSCRDT-SHKSVTTFQFAPQT 994
 QY 908 DNEPEIATMNNNDITLIERESFVDQREIEKADMLSEDSVVEBPQDQLESLOGKD 967
 Db 995 DKPCTIATKLAEDDLKVAHVHVAQLEGEV-----QIINSQ-----GKK 1035
 QY 968 DYFGSGOKEGEGFQKPIPASS-MPOGIGIPSSITQVYHAGRRFIVSPVESLRESK 1026
 Db 1036 GVGTKLATVDLPNGETPPTTAVMKSSAATASNTK-----PKIEIEKTPPT--RDAS 1087
 QY 1027 VFBEITIDYVAASQSPGMNLSHSSASLSLQAFSELRAQMTGPN-----TAPNPNH 1082
 Db 1088 QEPNNVQVTVWRKVSQ-----SNAESVQSIIPRPGIIVMSPTNQTD 1129
 QY 1083 TGPFPPVPPPLSSIAQVPTTAATAVPATSSPNDISTSVIOSEVTPTEBGIAGVAT 1142
 Db 1130 SAP-----PPGGAANK----- 1141
 QY 1143 STGVVTSGLPPIPVSEPLVSSVITIPAVYISTSPLOVPS-TSEIYVSSDAL 1201
 Db 1142 SRFQVTKSADPT-----ATPISSTISRTATVPIVAA-----TPNITISEPVI 1183
 QY 1202 YPSVYVYSSASAGSTATPGRPAVVSQAAGSTTGATLTGTSVTTSPFTASQOLST 1261
 Db 1184 -----VQPIYACVITHTLAPSPVSHS-LSGNSSPATTHSNMSISQISVSP----- 1229
 QY 1262 QSSSTSTPLAETVYVSAHSLDKTSHSTTGLAFSLAPSSSSSPGAVSSYISQPGGL 1321
 Db 1230 -----GRRFTVQGVSAQES-----GISSISIP----- 1252
 QY 1322 HPLVIPSIVASTPILPQAAGPTSTPLLPVQVPSIPPLVQVAVNPAVQOTLHNSQOPALJL 1381

Db 1253 HPEPTPAI-----TSCP--PPVSVPPV-----VSNQTLMEVAPKOT 1288
 QY 1382 PNOPTHCEVNDSTQPKAPGIDDIKYLEKRLSFEHSSSGAQAHSYSLTSLVIBST 1441
 Db 1289 PSATQNV-----DTQ-----HSSSTATATLIVSTPATVH-- 1319
 QY 1442 VTP-GIPTTAAP-----SKLITSTSTCLPPTMLPIGTVALPVTPVTPQVSTPVST 1494
 Db 1320 VTPISVPAVQEPVLVIDHSDVLTQ-----LDSELKRVSGVSHSPSTVES 1367
 QY 1495 TTSGVKEGTAP-----SKPPLTKAPVLPVGTETLPAGTLPSRQLPFPFG-----PELTOSQ 1544
 Db 1368 LTS-MTQPTIPLACQTVASIGQAPAV-IAAHAAHLIPNASVPOSFRLDAETGLAGLH 1425
 QY 1545 QPLEDL--DQRLRTISPEKITTVSANG-----PVSMAAPTA----- 1579
 Db 1426 EKLEALMKEODRREDMDDAIGITTTGDKDEIPIDTLKGIAEALGKVIHADGRETTPMP 1485
 QY 1580 -----ITEAGTOPQKGVQVKEGPVLTATSSGAGVPMKGRFQVSVAADGAQKEGKNKSEDAK 1635
 Db 1486 DHPDLTASTQO--LISPNDVLTMSA-----VEGASSTYME 1524
 QY 1636 SVTFESTSSSVLSSSPSTLVKPE-PNGITIPGISDVPSPAKHTTASEAKSDTQOP 1694
 Db 1525 DIDASTAVDASMMNSMPPGAQNSTDQIPAAAT--LSMD-QEQAQSTSSITRNTTG-- 1578
 QY 1695 TKVGRFOVTTANKVGRFVSKEDEKITDTKKGPAVASPEPMDLEQAVLPAVLPKKEPE 1754
 Db 1579 TKIATP-----NLFTALSTLCTHNRQP-- 1602
 QY 1755 LEPHSHNGSPSDPEA-----AFLSRDVDDSGS-----PHSPHQLSSKLS 1797
 Db 1603 -----NAPSRDETTAQTSPFNERIGGGGGATSFISGTPPSHP-----FPV 1648
 QY 1798 QMLSQSLNSFNSSYWSNDSIEDIEDLKLRLRLDKHLEKDIDQSRKHETSLT 1857
 Db 1649 SECDYDLK-----GQMDLESD--PEVICMIVHREOKHLEKQVELETSLRS 1695
 QY 1858 KLGKVPVAVIIPPAAPLGRRRRPTKSKSKSSRSSSLGNKSPQSLGNSQSAASVLHP 1917
 Db 1696 KI-RVPRATVNP--EMIGDDEADT--TLTALQALGNASLSLPAS----- 1736
 QY 1918 QOTLHPGNIPESGQNCLOPLKSPSSDNLYSARTSGAISVPSLSAPQCTSTNTVG 1977
 Db 1737 -----PPNETTETKNTVIVP-----SDVLAIRMTM-----SQSTKSSNVS 1773
 QY 1978 ATVNSQAQAQOP 1990
 Db 1774 VSSRRDNQSNAP 1786

RESULT 2
 S49611
 Probable serine/threonine-specific protein kinase pKpa (EC 2.7.1.-) - Phycomyces blakeslee
 C:Species: Phycomyces blakesleeans
 C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
 C:Accession: S49611; S59578
 R:Ruiz-Perez, V.; Murillo, F.; Torres-Martinez, S.
 submitted to the EMBL data library, November 1994
 A:Description: A novel serine/threonine protein kinase in the fungus Phycomyces blakeslee
 A:Reference number: S49611
 A:Accession: S49611
 A:Molecule type: DNA
 A:Residues: 1-633 <RUI>
 A:Cross-references: EMBL:Z46636
 R:Ruiz-Perez, V.L.; Murillo, F.J.; Torres-Martinez, S.
 Curr. Genet. 28, 309-316, 1995
 A:Title: pKpa, a novel Phycomyces blakesleeans serine/threonine protein kinase.
 A:Reference number: S59578; MUID:196120859; PMID:8590476
 A:Accession: S59578
 A:Molecule type: DNA
 A:Residues: 1-9,39-346 <RUI>

A:Cross-references: EMBL:Z46636
 A:Note: only part of the coding region is given
 C:Genetics:
 A:Gene: PKPA
 A:Introns: 60/1; 248/3; 277/2; 339/3; 534/1
 C:Superfamily: Phycomyces blakeleanus probable serine/threonine-specific protein kinase
 C:Keywords: phosphotransferase; protein kinase
 F:34-289/Domain: protein kinase homology <KIN>

Query Match 7.1%; Score 765.5; DB 1; Length 633;
 Best Local Similarity 35.2%; Pred. No. 5.5e-16;
 Matches 192; Conservative 91; Mismatches 165; Indels 77; Gaps 22;

```

QY 187 GGSAAKEQERSQOD--DIELETKAVGMSNDGRFLKFDIEIGSGFKTVYKGLDTETT 244
DB 5 GGVMMNPMDEVDSDNMDYE---KYLEASGNRGYSKLTNVLKGAIVKYYKAIIDBEA 60
QY 245 VEVAMCELDORRL--SESRQFEKE---EAMKLGLOHPNIVRYFDSWESIVKAKCIVL 299
DB 61 I-----NDNEITNVKVTROEFKDLGHEIDILKSVHPNITPHDAMYETE----FVF 109
QY 300 VTEIMTSGTLKTYLKRFRKM-KIKVLRSGROLKGLQFLHTRPTPIIHRDLKCDNIFIT 358
DB 110 ITLEMTSGTLREYIRKLTPLPNIKIKVRCRQILKGLAYLHGHPPIIHRDLKCDNIFIN 169
QY 359 GPTGVSXIGDLGLATLKRAAPKSVICTPEPAPEMEYEEK-YDESVDVYAFGMCLEMAT 417
DB 170 GAHGEIKIGMGALEMNGK-KYTVIGTPEFMAPEMEYEGYNEKVDIYAFGMCLEMAT 228
QY 418 SEVYSESCQAAQIVRTSGVKSAPDKVAIPEVKIIEGCIKONDERYSIKDLINHA 477
DB 229 GEYFPGECTYAVGVFKKVTITRPECISRVQDEPLLTVNICTLP-EDEMTKQELIEHR 287
QY 478 FFOETGVRVELAEEDGEKIAIKMLRIDIKKQKYNDAIEFSPFLERDVEDVA 537
DB 288 FLAVER--EYVLVSKMTMTL---LTLQV---VFKG--MDKLSVKEFPAADDTTAAV 336
QY 538 QENVESGYV-----CE-----GDHKTMAKIKDVSILIKRRE 570
DB 337 AEMIEQVLANCYOOLITCEINRLDIARNQGPDPKGEDEKIVMRENDIRSELRAKK 396
QY 571 QROLVRE---EQRKKKQ--EESLKOQVBSASQICIKLPASGICPIATTSASVST 625
DB 397 DLALAVRFEAKKCELEKHNI---LAEERCKET-IFALEQAKQIPLDLPQPOPQ 452
QY 626 QVEPE---EPEADQHQOLQYQPSISVLSDGTVDSCGSSVFESHVSQQTVSYSQHR 682
DB 453 QPQPOPQPOPQQLQQLQQLSLPQSTTSRPTSDDNSTNSTMSS-LSESLSKLVCVSGDE 511
QY 683 QAHST 687
DB 512 QVETT 516

```

RESULT 3

T06716
 hypothetical protein T29H11.220 - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #ext_change 17-Mar-2000
 C:Accession: T06716
 R:Querier: F.; Choiane, N.; Robert, C.; Broctier, P.; Wincker, P.; Cactolico, L.; Artigou
 submitted to the Protein Sequence Database, April 1999
 A:Reference number: Z15793
 A:Accession: T06716
 A:Molecule type: DNA
 A:Residues: 1-516 <QUB>
 A:Cross-references: EMBL:AL049659; GSPDB:GN00061; ATSP:T29H11.220
 A:Experimental source: cultivar Columbia; BAC clone T29H11
 C:Genetics:
 A:Gene: ATSP:T29H11.220
 A:Map position: 3
 A:Introns: 23/3; 36/2; 112/2; 187/1; 239/3; 295/1; 338/1
 C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homol

Query Match 6.9%; Score 745.5; DB 2; Length 516;
 Best Local Similarity 41.9%; Pred. No. 1.7e-15;
 Matches 169; Conservative 67; Mismatches 112; Indels 55; Gaps 12;

```

QY 194 QERSQOQODDIELETKAVGMSNDGRFLKFDIEIGSGFKTVYKGLDTETTVAAM--CE 251
DB 3 QDENNSSEEFVE-----IDPTGRGRYKREYVKGAKFKEVYAFQDLSEIVAMQVK 54
QY 252 LQDRKLTKESEKQFEKAEMLKGLQHPNIVRYFDSW---ESTVKGKKCIYLVTEIMTSG 307
DB 55 LDDKFCSEEDLDRLYSEVHLKTLKHKSITKFTYSWIDHQMTIN-----LIREVFTSG 108
QY 308 TLKTYLKRFRMVKIKVLRSGROLKGLQFLHTRPTPIIHRDLKCDNIFITGPTGSYKIG 367
DB 109 NLRQYRKHKCVDRALKKWSRQTLSEGLVYLHSDPVIHRDLKCDNIFITNGNGEYKIG 168
QY 368 DLGLAT-LKRAAPKSVICTPEPAPEMEYEEKYDESVDVYAFGMCLEMATSEVYSECO 426
DB 169 DLGLAALIKRAASHSVICTPEPAPEMELYEDYNVLVDIYAFGMCLELVTFEYPSCT 228
QY 427 NAAQIVRTSGVKSAPDKVAIPEVKIIEGCIKONDERYSIKDLINHA--FOET 483
DB 229 NAAQIVRTSGVKSAPDKVAIPEVKIIEGCIKONDERYSIKDLINHA--FOET 287
QY 484 -----GVRELAEDDG-----EKIAIKMLRIDIKKQKYNDAIEFSPFLERDVEDVA 520
DB 288 ENVSHKENGNGNGVYDKLSDSEVGLTVYEGQRKDLNITFLKRLTD---SKQIRN-- 342
QY 521 AIEFSPFLERDVEDVAQAEVWS-GYVCEGDHKTMAKIKDVSILIKRRE 562
DB 343 -IHFPNIETDTSFSAIENVEBELDLDDODISTIAKMDITEI 384

```

RESULT 4

T46059
 MAP kinase - Arabidopsis thaliana
 N:Alternate names: protein T18N14.10
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #ext_change 04-Feb-2000
 C:Accession: T46059
 R:Deleeny, M.; Berger, C.; Cooke, R.; Grellert, F.; Laudie, M.; Mewes, H.W.; Lemcke, K.;
 submitted to the Protein Sequence Database, December 1999
 A:Reference number: Z23013
 A:Accession: T46059
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-547
 A:Cross-references: EMBL:AL132968
 A:Experimental source: cultivar Columbia; BAC clone T18N14
 C:Genetics:
 A:Map position: 3
 A:Introns: 24/3; 113/2; 188/1; 240/3; 350/1
 A:Note: T18N14.10

Query Match 6.5%; Score 702.5; DB 2; Length 547;
 Best Local Similarity 31.9%; Pred. No. 3.3e-14;
 Matches 189; Conservative 88; Mismatches 202; Indels 113; Gaps 17;

```

QY 195 EERSQOQODDIELETKAVGMSNDGRFLKFDIEIGSGFKTVYKGLDTETTVAAMCELOD 254
DB 2 EISSASDDSIAYVET-----DPSGRYGRFREVILGKAMKTVYKAPDQVGMVAMNQVYL 56
QY 255 RKLTKSER--QRFKEAEMLKGLQHPNIVRYFDSWESIVKAKCIVLVTEIMTSGTLTKY 312
DB 57 NEVFRSPPELQLYSEVHLKMLNHSIIRYCTSWIDV--NRRTNFITELFTSGTLREY 114
QY 313 LKRFRMVKIKVLRSGROLKGLQFLHTRPTPIIHRDLKCDNIFITGPTGSYKIGLGLA 372
DB 115 RRYGVKIVIRAKISAKQQLNGLAYLHGHPPIIHRDLKCDNIFVNGHGLGVYKIGDLGLA 174
QY 373 TLKRAAS-PAKSVICTPEPAPEMEYEEKYDESVDVYAFGMCLEMATSEVYSECOAAQI 431

```

Db 175 AILRGSQNAHVSITPEPMAPPELYEDVNLDIYSPGMCULEMLTGEYPSRCTNPACI 234
 432 YRRVTSVGPASFDKVAPEVKEIIEGCIRQNKDERYSIKDLNMAFQOETGVVELAE 491
 235 YKVTSSGKLPDSFHLIQTHTAQRFGVKCL-ETVSRLPAKELLADPFL-----AATDE 286
 492 EDDGEKIAIKMLRIEDI-----KKLKGKKYKME-----520
 287 RDLAPLFRLLPQOAIQNLANGTVNHLBPTDPTRTTMDKSTIGCKNSEDHITFLOVOLI 346
 521 -----AIEFSPDLERDVEDVAQEMVESGVCEGDKHTAKAIDKRVAS-LIKRRRQR 572
 347 DGDGHRNRIQFPFNITLSDTPELVALEMVKELETMDPLEIAMINENISLIVPMNRAND 406
 573 QLVREBQEKKKOEBSLKOQVSSASQGTGKQLPSASTGIPASTTSASVSTQVEEPP 632
 407 SSRHRE-----SFGHDEDEDNGTDEGRTRL-----FSSASSSHDSPAVAREN 448
 633 EADQHQOLQYQOQPSISVLSGTVDSGQSSVFTESRVSSQOTVSYG-----SQHQAHST 687
 449 NDS-----SNDVIPD--MDDGNRSS-----NRLNLSSTHYSPALDDDONQOORRR 493
 688 GTVPGHIPSTVQAQSQ-----PHGVPPSSVAQSQSQSQSQSS 725
 494 VRLQQRMSLVDTRTQVLRHSIMELINKRRGRGFDPNTN-----ELQOPQSSST 541

RESULT 5

TS1099
 MAP kinase [imported] - Arabidopsis thaliana
 C/Species: Arabidopsis thaliana (mouse-ear cress)
 C/Date: 21-Jul-2000 #sequence_revision 21-Jul-2000 #text_change 21-Jul-2000
 C/Accession: TS1099
 R/Coke, R.M.; Berger, C.; Delzeny, M.
 submitted to the EMBL Data Library, May 1999
 A/Description: Analysis of Arabidopsis thaliana gene structure by cognate cDNA sequencing
 A/Reference number: 225297
 A/Accession: TS1099
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: mRNA
 A/Residues: 1-549 <COO>
 A/Cross-references: EMBL:AJ238802; PIDD:CA43520.1

Query Match 6.5%; Score 702.5; DB 2; Length 549;
 Best Local Similarity 31.9%; Pred. No. 3.3e-14;
 Matches 189; Conservative 88; Mismatches 202; Indels 113; Gaps 17;
 195 EERSQOQDDIEBELTKAVGMSNDGRFLKPIEIGRSFKTVYKGLDTEETVEVAMCELOD 254
 4 EISSASDSDSLAVET-----DPSGRYGRFREVLOKGMKTVYKAFDVLGMEVAMNQVKL 58
 255 RKLTKSER--ORFKEAEMLKGLQHPNIVRPFYDSWESTVKGKKCIIVLTELMTSGTLKTY 312
 59 NEVRSRPEPQLRTYSEVHLKLNLMHESIIRYCTSMIDV--NRRTFNFITELFTSGTLREY 116
 313 LKRRKWKIKYLRKWCQIILKGLQPLHTPTPIIHRDLKCNITPTGTSVKXIGDGLA 372
 117 RRRKQXVDIRAIKSWAQIILGLAYLGHDPVPIHRDLKCNITFVNGLGVKIGDGLA 176
 373 TLKRAS--FAKSVITGPEPMAPPELYEDVNLDIYSPGMCULEMLTGEYPSRCTNPACI 431
 177 AILRGSQNAHVSITPEPMAPPELYEDVNLDIYSPGMCULEMLTGEYPSRCTNPACI 236
 432 YRRVTSVGPASFDKVAPEVKEIIEGCIRQNKDERYSIKDLNMAFQOETGVVELAE 491
 237 YKVTSSGKLPDSFHLIQTHTAQRFGVKCL-ETVSRLPAKELLADPFL-----AATDE 288
 492 EDDGEKIAIKMLRIEDI-----KKLKGKKYKME-----520
 289 RDLAPLFRLLPQOAIQNLANGTVNHLBPTDPTRTTMDKSTIGCKNSEDHITFLOVOLI 348
 521 -----AIEFSPDLERDVEDVAQEMVESGVCEGDKHTAKAIDKRVAS-LIKRRRQR 572

Db 349 DGDGHRNRIQFPFNITLSDTPELVALEMVKELETMDPLEIAMINENISLIVPMNRAND 408
 573 QLVREBQEKKKOEBSLKOQVSSASQGTGKQLPSASTGIPASTTSASVSTQVEEPP 632
 409 SSRHRE-----SFGHDEDEDNGTDEGRTRL-----FSSASSSHDSPAVAREN 450
 633 EADQHQOLQYQOQPSISVLSGTVDSGQSSVFTESRVSSQOTVSYG-----SQHQAHST 687
 451 NDS-----SNDVIPD--MDDGNRSS-----NRLNLSSTHYSPALDDDONQOORRR 495
 688 GTVPGHIPSTVQAQSQ-----PHGVPPSSVAQSQSQSQSQSS 725
 496 VRLQQRMSLVDTRTQVLRHSIMELINKRRGRGFDPNTN-----ELQOPQSSST 543

RESULT 6

T02951
 Probable mitogen activated protein kinase - rice
 C/Species: Oryza sativa (rice)
 C/Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 24-Mar-1999
 C/Accession: T02951
 R/Kim, C.Y.; Cheon, S.Y.; Cho, M.J.
 submitted to the EMBL Data Library, July 1998
 A/Description: Identification and characterization of fungal elicitor responsive rice ger
 A/Reference number: Z14788
 A/Accession: T02951
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: mRNA
 A/Residues: 1-677 <KIM>
 A/Cross-references: EMBL:AF080436; NID:G3450841; PID:G3450842
 A/Experimental source: cultivar Mallyang 117
 A/Genes: MEK1

Query Match 6.3%; Score 678.5; DB 2; Length 677;
 Best Local Similarity 32.4%; Pred. No. 2.1e-13;
 Matches 176; Conservative 75; Mismatches 186; Indels 107; Gaps 14;
 218 GRFLKDDIEIGRSFKTVYKGLDTEETVEVAM--CELODKLTKSERORFKEAEMLKGL 275
 23 GRYGRINDVIGKASKTVYKAFPEYQMEVAMNQVKLHDFLOQBEDERLYCEIHLIKTL 82
 276 QHPNIVRPFYDSWESTVKGKKCIIVLTELMTSGTLKTYLKRFXWKIKYLRKWCQIILKGL 335
 83 KHNIMKFTYSWDV--SRNINPITFMTFSGTLRQKQHRVNIWAVKWCQIISGL 140
 336 QPLHTPTPIIHRDLKCNITFVNGNQEVKIGDGLAAILRKSAAVHCVTPEPMABEV 395
 141 LYLSHDPPIIHRDLKCNITFVNGNQEVKIGDGLAAILRKSAAVHCVTPEPMABEV 200
 396 EEKYDESVDVYAFQMCULEMATSEYPSSECONAAQIYRRVTSVGPASFDKVAPEVKEI 455
 201 EESTNELVDIYSGMCULEMATSEYPSSECTHPQIYKTIISGTRPALVKVDPVNRQF 260
 456 IEGCIRQNK-----DERYSIKDLNMAFQO-----ET 483
 261 VEKLTATASRLSARVILKDPFLQVDLVFCPDGNGSNLMNYLRQPLYOHAYSTVSMNSN 320
 484 GVRVEL-----AEDGEKIAIKMLRIED-----IKLKG-----514
 321 GLSESIDSPTEDEDMCEDDIDAGDILFNGHEDEPLGNVDITTIKRSSEDSITFLRL 380
 515 KYKQNEA-----IEFSPDLERDVEDVAQEMVESGVCEGDKHTAKAIDKRVSLIKRRR 570
 381 RIANDGHRNRIYFPFIHADTALSVALEMDIDIDHEVTRAEMIDGVSALVDMR 440
 571 QRLVREBQEKKKOEBSLKOQVSSASQGTG--IKQLPSASTGIPASTTS--ASVSTQ 626
 441 PGGIESQDTTYGNCG-----SNVSSCGSLYAVMCAARGCHGADLHGRFEDITFQ 493
 627 VEPEPEADHQOLQYQOQPSISVLSGTVDSGQSSVFTESRVSSQOTVSYGSGHQAHST 686
 494 ANGEQTDLOD-----SGGSSDDGGQOT--QHVXQDEAV-----HS 526

QY 687 TGTV 690
Db 527 NGFV 530

RESULT 7
T34434
hypothetical protein K06A9.1a - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 18-Feb-2000
C:Accession: T34434
R:Geisel, C.; Gattung, S.
Submitted to the EMBL Data Library, December 1996
A:Description: The sequence of C. elegans cosmid K06A9.
A:Reference number: 221525
A:Accession: T34434
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-2232 <GET>
A:Cross-references: EMBL:U08046; PIDN:AACT0890.1; GSPDB:GN00028; CESP:K06A9.1a
A:Experimental source: strain Bristol N2; clone K06A9
C:Genetics:
A:Gene: CESP:K06A9.1a
A:Map position: X
A:Functions: 38/1; 75/3; 103/3; 132/2; 158/2; 222/1; 1088/1; 1367/1; 2039/1; 2049/1; 2075/1

Query Match 5.5%; Score 591.5; DB 2; Length 2232;
Best Local Similarity 22.8%; Pred. No. 2.5e-10;
Matches 385; Conservative 211; Mismatches 614; Indels 481; Gaps 72;

QY 587 SSIAQGVESASQGTG-----IKQLPSATGIPATSTASVSTQVEPEPEADQHQ 639
Db 593 SSSPQSSQSPAPENTSTTPSQTSQSPSPSMN-PSSSTPTGSSQSTIPEG----- 643

QY 640 LQYQGSISVLSGCTYDSGGSSVFTEKRVSSQQTYSYSGHQAHTGTVPCHITSTVQ 699
Db 644 -----STASSPTSGTSTFSVATE-VTSQSTVPSGS-----SLGT-----QSTN 681

QY 700 AAGQPHGVVPPSS-VAQGSQSQGPSSSLTGVSS--SOPLOHPQ----- 741
Db 682 SSSPSSSLSTSGMTLTNSBPSSSTQSGAOSTITTPSPNSQSTSLSSSTGATTS 741

QY 742 --QGIQCTAPQ-QTVQVSLQSTSSSEATTAQPVSQ--PQAPVLPVVSAG--KQSTQ 794
Db 742 SCSAGTMTSPSSQSSSVSGSSTSPASATSGEMTSQSGTQTPGSSVSTSAALINSTQ 801

QY 795 GVSQVAPAEVNAQQAQOPTLLASSVDSAHSDVASGMSDGNENVPSSSGRHEGRTYK 854
Db 802 SVSTNSPGSTV-----TRPSTVSGSTSSG-STVVGSTEAISTGSSVASSSPAPSTQ 853

QY 855 HYKSVRS-----RSRHEKTSRPLRLINLNSNKGDRVVECOLETHNRMKVTFKPLDD 908
Db 854 NRPSTSSSSSMTQTPYPSQSTSPVSSSTTPSGPGTTLTSTSPSQSTTGSTQCS 913

QY 909 NPEIATIVNNDFLAIERESFVDQREIIEKADMLSDVSVBEPDQGLSLQK-- 966
Db 914 TSPGIST-----TSEMTSQG-STQPGSTGTVTQPSV 947

QY 967 DUYGFGSQKLEGEFK--QPIPASM-----PQ----- 992
Db 948 SOSTSSGTVTGSTEGSSSPSTQNTNPSTSSGSSMSTQTPQSSQSTSPVESSTGA 1007

QY 993 --QIGIPTSILQVHVSAGRFRIVSPVPSRLRESKVPFSEITDVTVAASAGPG----- 1045
Db 1008 TSSSGSPGTLTNI-----SPSPSPSTTSSSGSSTSPVSTTSSQSTETPGTGT 1059

QY 1046 ---NLSHSAGSLSLQQAFLSELRRAQMTGEPVADPNFSH-----TGP 1086
Db 1060 VTKPSTVSGSASGSAITWMS--TEASSTGSGSTSPNPSQSTSPSTGATSPSGSGTT 1117

QY 1087 FPIVPPFLSLAGVPTTAATAVPAVATSSPPNDI-----STSVIQSEVTVPTREGIAGVAT 1142

Db 1118 LTISSPSSQSSSTIGSSQGSTSPVSTTS--GDMTSQSTQIPIGSTGTVTQESTGSGST 1175

QY 1143 ST-GVATSGG---LPIPVSESPVLSVSVSITIPAVVSISTTSPSLQVPTNSSEIVSS 1198

Db 1176 STSGEITTSQSTQTPPSSSLTSPALSTQO-----SVSTNSP-----GS 1215

QY 1199 TALPSEVTSANASASAGSTATPGPPAVVSSQAAGSTVVG-ATLTVSTTSPSTAS 1257

Db 1216 TVTQPTVAGSTSS---GSTVITG-----STGSSSTSSGSSATSLSSSPVSTSQ 1263

QY 1258 QUSIQUS--SSTPTPLAEV--VSAHSLDKTSHST--TGLAFSLADPS--SSPG 1308

Db 1264 SPNPSTSGSTTPPNSQSTSPVSTTGEMTHSGTQTPSTIGSTVTCSTVSGNSG 1323

QY 1309 AGV-----SSYIQPGGLHPVLVPSVASTPILPQAAGP-----STPILP 1349

Db 1324 STVTTSSSNASSTSGSFKTSPSSISVPNPSSPITTPASSTSGSTSDVSVSTSLAP 1383

QY 1350 QVPSIPPLVQPVANVAVOQTLIHSQOPALLPNCPTHCEVSDTQPKAPGIDIKTL 1409

Db 1384 LSSSLP-----STVPSSTQSPSTSESSKASSP--VPQSTSTPTNPICSTESSL 1434

QY 1410 EKLKSLFSEHSS--SGAQAIVSLSTSLVIESTVTPGIFTTAVAPSKLLTSTTCLPT 1468

Db 1435 LSTTISGSTQHTTMSKASSGSTSPNSQSTGTVTWSGST-----SGVSTSSASTQPM 1490

QY 1469 NPLGTVALPVTPVTPRGVSTPVSTTSGVKGTPSKRPPLKAVLPVGT--ELPAGT 1526

Db 1491 STSQSSA-----GSTVASTASPAASSTPSS-----TGTMSSTSGT 1529

QY 1527 LPSEQLPPEPGSLTOSQOPLDLDAQLRRLTSPXKIT-VTSAVGVSMAPATPAITAGT 1585

Db 1530 V-----GSLTSSS-----TTMSASQSTSTVTGSSSTSGVSTSSAST 1569

QY 1586 QPKGVSYQKEGPVLTATSSGAGVFKKGRQVSVADGAKE--GNKKEDAKSVFESS 1642

Db 1570 QPMSTSQSSASGTVASTAGIVSTVPSSTGTGWSSTGTVSGSTTSSSTTASASQ 1629

QY 1643 T-----SESSVLSSSSESTIVKPEPNGLTIPIS-----SDVPEE- 1678

Db 1630 TSGTVMGSSSTSGVSTSSAST--QPMSTSQSSASGTVASTTGLVSTVPSST 1685

QY 1679 -----AHKTJASEAKSDTGOPTKVRFGQ--YTTANKVGRFESKTEBK 1720

Db 1686 GTMGSTSGTVSGTSSSTASASQSTGTVTGMSSSTSGVSTSSASGQPMGTSQCS 1745

QY 1721 ITDTKKEGVASPPFMDLEQAVLPVITPKKEKPELSEPHLNGPSSDPPAALSRDVDG 1780

Db 1746 SAGSTVASTASP-----AASSTAP-----SSTGTMGSTSGTVSGTSSQSTAA 1790

QY 1781 SGPSPH--SPHOLSKSLPQNLQSLNSGFSNYSMSDNEIDEDDLKLELRRLDKHL 1838

Db 1791 STTSHGTVTLTSSSTSSNQSTQSSSVSTVASS----- 1827

QY 1839 KEIODLQSKHEISLTYKLGVPPAVIIPPAAPLSGRRRPTKSGSKSSRSSLAGNK 1898

Db 1828 -----TAGLVERST-VPSTGTMGSTSGT--VGSITSESTTAA 1865

QY 1899 SPQL-----SGNISGQSAASVLAHQDTLHPGNIPESSGQNLQPLKPPSSDNLVYA 1951

Db 1866 SSQTSVTVMGSSSTSGVSTSSASTQ-----PQMSQO-----GSSAGSTVA 1908

QY 1952 FTSDGALS--VPSLSAPQGSTSTNTVGATVNSQAQOPAPMTSRKGTPTDLHKLV 2008

Db 1909 SSTAGIVSTVPS-STGMGSTSGTVSTI--SESTIASSTSSQGS----- 1954

QY 2009 DNMADAMLSGRRSKGMHNYEGPMARKFS-----APQQLCISMTSNLGSAP--- 2058

Db 1955 -----TVTIGSTGT-----NPSSPRLSQITTPPSPGSTSTQSLSPSSSPST 2001

QY 2059 --ISAASATSL 2067

Db 2002 HSSVSSSEGTMM 2012

RESULT 8
T45025
muscin MUC5B, tracheobronchial (imported) - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 21-Jul-2000
C:Accession: T45025
R:Gessy, J.L., Guyonnet-Duperrat, V., Porchet, N., Aubert, J.P., Laune, A.
J. Biol. Chem. 272, 3168-3178, 1997
A:Title: Human muscin gene MUC5B, the 10.7 kb large central exon encodes various alternat
A:Reference number: Z22899; MUID:9716151; PMID:9013550
A:Accession: T45025
A:Status: preliminary; translated from GB/EMBL/DBD
A:Molecule type: DNA
A:Residues: 1-3570 <DES>
A:Cross-references: EMBL:Z72456; NID:91834502; PIDN:CA96577.1; PID:91834503
A:Experimental source: placenta
C:Genetics:
A:Gene: MUC5B

Query Match 4.6%; Score 494; DB 2; Length 3570;
Best Local Similarity 19.4%; Pred. No. 3e-07;
Matches 419; Conservative 226; Mismatches 705; Indels 814; Gaps 88;

QY 596 SSASQGIKOLPASTG-----IPTASTSASVTOVEPEPEADQOQLOQYQOPSISV 650
DB 1177 TTYTANSVPIPISSSLGTTWRLSQTTTPATMSTAPSTPETAH-----TSTVL 1227
QY 651 SDGTVDSGQSSSVFTBS-----RVSSQOTVSYSQCHEQAH-----STGTVP 691
DB 1228 TATATTTAGTAGVATPSTGTATHTKVPTTTTGTATPSSPGALTPVWISTTTP 1287
QY 692 GHTPSTVVOAQSOPHGVY-----PPSSVACG-----QSQGQPS--SSLTVSSS 733
DB 1288 TTRGSTVTPESIGTHTATVLTITTTTATVAGSMATPSSSTQTSQTPPSLTITATITAT 1347
QY 734 OPTIHOQOQGIQOTAPPOQTVQYSLQSTSEATTAQVSOQAQVLPVQVSAKOST 793
DB 1348 GSTTNSSTGTGTPPIPVLTIT--ATTPTATSSVTPSSALGTHHPV--PNTTA--TT 1401
QY 794 QGVSAVAPAEVAV-----AQPOATPPTLIASSVDSA-----HSDVASG-- 832
DB 1402 HGRS--LSPSSPHVTRATWATSATGTLGTHITPSTGSHTPATGTOHSTPALSPH 1460
QY 833 -MDDGHNVSSSGRHGR--TKRHKVSRSR-----HEKSRPLKILNS-- 880
DB 1461 PSSRTTESPPSPGTTTGHATTATRTTATATPSKTRTSTLLPSSPTSAPIITTVTMGCEP 1520
QY 881 -----NKGD-----RV 886
DB 1521 QCAMSEMLDYSYMPGSGSDPDTYSNIRAGAVCEQPIGLECRAQAOQVPLRELQY 1580
QY 887 VECQLE---THNRKV--TFKEDLD-----GDNPEIATIMVNDPILAIRESF 931
DB 1581 VECSLPFGALVCRRREQYQKFCMCFNVEIRVCCNVGHCFTPAT-- 1624
QY 932 VDOVREIKADMLSEADVVEPEBDOG-----LESIQGKDYGFSQKLEGEFKQPIP 986
DB 1625 -----SSTATPSTPGTWTWILTEQTTATTTATGTSSTPISDTAP 1667
QY 987 ASS-MEQOIGIPTSLSLQVHSAGRPIV-----SPVRESL----- 1022
DB 1668 PRKVLVSQATTPATSKATSSSPKATLLPVLTSTATSTATSPFPISSSLGTTGTS 1727
QY 1023 -----RESKVPSEITDVAASTAGSPGMNLSHSASSLQQA-----FSELR 1066
DB 1728 ONRPHPMATSTIHPSSTPEITHTSTVLTKATTTATSSMSTPSTPGTWTLELT 1787
QY 1067 AQMTBG--PNTAPNNSHTGPTFPVPPPLSSIAQVPTTAAATAPVATSPSPNDISTSV 1124
DB 1788 AATTALPHGTPS--STPGTWILPESTTATVPTGSTATAS----- 1830

QY 1125 IOSEVTPTEBGIAGVATNGVTSGLPIPIPVSEPVLSVVSSTIPAVSIS----- 1179
DB 1831 -STRATAGTLKVLSTATTATPVISSRATPSSSPGATLALPALMSATATTPATSVTAIPS 1889
QY 1180 -----TTSF-----SLQVPTSEIIVSSTALYPSVTVSATASAGSTAPGPK 1224
DB 1890 SLGTATRLSQTTTPATMSTAPSTSPETVHTSTVLTTATTTTRT-----GSVAIPSS 1944
QY 1225 PRAVVSQAAGSTTVATLTSTVSTTSSFPSTASQSLQSSSISTPLAETVVSASHLD 1284
DB 1945 PGTATHTKVPITTTTGTFT--ATPSSSPGATLTPVWISTITTPTRGSTVPS--SIP 1998
QY 1285 KTSN-----SSTGLAFSLASPPSSSSPGAGVSSVTSQPGIHLPIVPSVASTPILP 1337
DB 1999 GITHATVLTITTTTIVATGSMATPSS-----TQTSQ-----TPSLITATTTI 2042
QY 1338 QAAGPTSTP-LIPQVSIPIPLVQVAVNPVAVQOTLI-----HSQOPP----- 1378
DB 2043 TATGTTNPSSTPGTPIPIPVLTATTPATATSSVTPSSALGTTTRPPVNTTATTHGR 2102
QY 1379 ALLPNQPH-----HCPVDS--DTPPKAFGIDIDITLEE 1411
DB 2103 SLPPSPHVTPTAMTATSGILGTHITPSTGSHTPATVGTTPSTP----- 2152
QY 1412 KLRSLFSEHSSGQAQASVSLFSLVIRESTVTPG-----IPTAVA--PSKILSTSTQCL 1465
DB 2153 ---ALSSPHPSKTTSPSS-----PGTTTGHGRGTSRTTATATPK--TRISTILL 2199
QY 1466 P--PYNLPLGTV----- 1475
DB 2200 PSSPTAPITTVVTTGCEQCAMSEMLDYSYMPGSGSDPDTYSNIRAGAVCEQPIG 2259
QY 1476 -----ALPVTVPVTPGOV-----ST 1490
DB 2260 LECRAQAPGVPLRELQYVVEGSLDFGLVCRRBQYKFCMCFNVEIRVCCNVGHCPST 2319
QY 1491 PV--STTTSQVKKGT-----APSKPLTKAPVLPVQ-----TELPAQ 1525
DB 2320 PATSSATPSSPGTWTWILTKLTATTTSTSTSTAPSTQGP--PAGPHVSTIATTP 2377
QY 1526 TLPSEQLPPPPGSLTQSQQPLEDLAQRLTSPKXITVTSAGVSM-----A 1575
DB 2378 TVTSSKATPSSPG--TATALP-----ALRSTATPTATSTFALPSSSLGTTWRLSQTT 2430
QY 1576 APTAIIEAGTQPKQSVQYKEGPVLTSSGAGVCFKMGKQVSAAGAOKENKSEDAK 1635
DB 2431 TPAATMSTAT--PSSTPEVHTSTVLTTA-----TTGAGAGVAATPSTGTAT 2478
QY 1636 SVHFESESTSES--SVLSSSSPE-----STLVKEPNGITIPGISDVPESAHKTTAS 1685
DB 2479 TTKVPTTTTGTFTVTPSSSPGTRTPFWIISTTTPTTSGSTV--TPSSIPGTHHP-- 2534
QY 1686 EAKSDTGOPTKGRF-----QV 1702
DB 2535 -VLTITTOVATGSMATPSSSTQTSQTPSLLITATTTATGSTTNPSSPTGTPPIPEL 2593
QY 1703 TTTANKVGRFVSYTEKDLITDKKEGPVASPPMD-----LEQVLP 1745
DB 2594 TTTAT-----TPATISTVPSALGTHHPVPPNTTATTHGSLSPSHVTRATWTS 2648
QY 1746 VIPKKEKPELSBPSHANGPSDEEAAFLSRDVDGSGSPHSQPLS----- 1791
DB 2649 TSGTLGTHITERS--TGTSHTPAA--TGTTTSPALSSPPSSRTTESPPSPGTTTP 2704
QY 1792 -----SKSLPQONLSQSL-----SNSFNSY--MSD 1816
DB 2705 GATTATRTTATATPKTRTSTLLPQOPSAPITTVTTGCEQCAMSEMLDYSYMPG 2764
QY 1817 NESDIE-----DEDKLELR-----LADKLEKI 1841
DB 2765 SGSDPDTYSNIRAGAVCEQPIGLECRATAQGVPLGELQYVEGSLDFGLVCRRBQY 2824
QY 1842 QDIQSRQKHEIBSLYTKLQVP--PAV----- 1866

Db 2825 GKFMCEHYEIRVCCNVCNHCSTPATSTAMPSTPGTWTILTELTATTATTAAGSTA 2884
Qy 1867 -----IIPPAALSGRRRRRTKSKGSKSSRSS--LGKSKPOLSGNLSCQSAAYLH- 1916
Db 2885 TPSTPGTAPRPVKLTSPATPT-ATSSKATSSSSPRATTLPLVLTSTATSTSTVTP 2943
Qy 1917 PQQLLHPGNIPESSGQNL-LQPLKPSSSDNLYSAFTSDGALSVSLAPGGTSTN 1974
Db 2944 PSSLTGTTGLPEQTTTPVATMTSTIHPSSEFETHTS-----TVLTATATATSTTS 2996
Qy 1975 TVGA-----TVNSQAQAQAPAMTSSKGTFTDDHLKLVNMAPDAMNLSG-R 2021
Db 2997 TPSTPGTWTILTELTATTATTAAGTGPATPSTPGT-----TWILTELTATT 3046
Qy 2022 RGSKGNHNYEGPMARKFSAPOGLCI-----SMTSNLGSAAPISASATS-LGHFTK 2072
Db 3047 TAGTG-----STATLSTPGTWTILTEPSTRTVTAAPGSTRATASTOATAGTPHVT 3099
Qy 2073 SMCPPQOYGFPAIPFGAOWSGTGAPAPQPLGQFQPVGTASLQNFENISLQKSIINPGSN 2132
Db 3100 TATTPVTSSKATPSSSPTATLPA-----LRVATTPATSTPTAIPSSS 3145
Qy 2133 LRTT 2136
Db 3146 LGTT 3149

RESULT 9

S69703
HKR1 protein precursor - yeast (Saccharomyces cerevisiae)
N/Alternate names: protein YDR420w
C/Species: Saccharomyces cerevisiae
C/Date: 20-Aug-1996 #sequence_revision 06-Sep-1996 #text_change 23-Mar-2001
C/Accession: S69703; A53382
R/Dietrich, F.S.
submitted to the EMBL Data Library, August 1995
A/Description: The sequence of S. cerevisiae lambda 3641 and cosmids 9461, 9831, and 941
A/Reference number: S69555
A/Accession: S69703
A/Molecule type: DNA
A/Residues: 1-1802 <DIE>
A/Cross-references: EMBL:U03007; NID:g927685; PIDN:AMB64857.1; PID:g927691; MIPS:YDR420w
R/Kasahara, S.; Yamada, H.; Mto, T.; Shiratori, Y.; Miyamoto, C.; Yabe, T.; Nakajima, T.
J. Bacteriol. 176, 1488-1499, 1994
A/File: Cloning of the Saccharomyces cerevisiae gene whose overexpression overcomes the
A/Reference number: A53382; NCID:94156857; PMID:8113191
A/Accession: A53382
A/Molecule type: DNA
A/Residues: 1-581, 'A', 583-593, 'A', 595-1802 <KAS>
A/Cross-references: EMBL:S69101; NID:g545659; PIDN:AMB30051.1; PID:g545660
A/Experimental source: YNN295
A/Note: sequence extracted from NCBI backbone (NCBIN:144410, NCBIP:144411)
C/Genetics:
A/Genes: SGD:HKR1
A/Cross-references: SGD:S0002828; MIPS:YDR420w
A/Map position: 4R
C/Keywords: calcium binding; glycoprotein; transmembrane protein
F/1-21/Domain: signal sequence #status predicted <Sig>
F/22-1802/Product: HKR1 protein #status predicted <MAN>
F/1483-1508/Domain: transmembrane #status predicted <TM>
F/1645-1656/Domain: calcium binding #status predicted <CAL>

Query Match 4.4%; Score 478; DB 2; Length 1802;

Best Local Similarity 22.6%; Pred. No. 4.3e-07; Mismatches 604; Indels 328; Gaps 64;

Matches 333; Conservative 209; Mismatches 604; Indels 328; Gaps 64;

Qy 596 SSASQGTGKQLBSAGTGTASTTASVSTOVEPPEPADQOOL--QYQOPSISVLSQ- 652
Db 29 STSYNNNGISTSTSYST-----SAISSTGSNNKENATSSSETTVMAGYGGESSTIME 83
Qy 653 -GTIVSGGSSVFTTERVSSQOQTVSGSHQAHSGTGVGHIPSTVQAQSGPHGV---- 707

Db 84 QETGTSQOYISVTTTQTSPTDMS-SVKSTEIAIPSSSI---VEPLQSVSEDSQISQTL 139
Qy 708 -YPPSSVAQOQSGQPS-SSSLTGVSQCPQIHPQOQGGIQQTPPQOQVQYSLISQTS 765
Db 140 SHNPKSYAESDSDTSTSESSSVITSTDSAVPREISPTITTDQSISKEGTLAQTSI 199
Qy 766 SEATTAQPVQAPQVLPQVLSAGKQSTQGVQVAPVAVQAPQATPTLASSV--- 822
Db 200 SEIT-----KLAQVNRVSC-----ISSITASTIDGSSSETO-TDESNVTSFE 243
Qy 823 DSAHSDVASGMDGNNENVPSSSGRHGRTTKRHYRKSVRSRSRHEKTSRPLRLIYNSK 882
Db 244 NSVEEYAMSKSQLSSESYSSSTYVSG-----GESADAKTSSSP-ITPSSS 289
Qy 883 GDVVEQGLETHNRKMYTFEFDLDGNDPELAIIMNNDITLIERSPVDQVEITEKA 942
Db 290 YSQTTSTSESSRVAA-----GVSRPSSITQTTSI----- 320
Qy 943 DEMISEDVSVPEPGDQGLESLQCKDPYFGSQKLEGEFKQPIPASMPQOI-----GI 996
Db 321 DSFSMSREVELSTY-----YDSAGVPPQDELIVDRPATSTETSEASQGV 367
Qy 997 PTSLTQVYHS-AGRRFVSPVESRLRESKVPSEITDTVAASTAQSPGMNLSHASL 1055
Db 368 SRESNTEFAVSSISTTFIVSASDT-----VSTSTNTVFPVSSVSHSTVHAATSSSTYI 421
Qy 1056 SLQQAFLRRAQMTGPNTPAPNPFSGTGFVVPPLS---SIAGVPTTAATAPVPA 1112
Db 422 S-----SSLVSPSLSASVSHFG-----VAFPPAYISFSVPVAVVST---Y 462
Qy 1113 TSSPPNDISTVQSEVYVTEEGIAGVATSTGVVNSGGLPIPVSESPSLSVSITI 1172
Db 463 TSSP-----SASVVP-----SAYSS-----PSVPA---VSTYTSPPSA 496
Qy 1173 PAVVSI-S-TTSPSLQVPTSTSELIVSS---TALYPSVTSATASAGSTATPGKPPA 1227
Db 497 PAIISSTYTSPPS-APVAVSSTYTSPPAPAIISSTYTSPPAPAVSSTYTSPPAPAI 554
Qy 1228 VVSQQAAGSTYVATILTSVSTTSPSTASQSLQSSSTPTTLAETVVSASHLDKTS 1287
Db 555 AISTYTSPPS-APVAVSSTYTSPPAPAIISSTYTSPPAPAVSSTYTSPPAPAI 612
Qy 1288 HSGTTG-----LAFSLAPSSSSPPGAGVSYISQPGALHPIVFPVIASTPLTPAACP 1342
Db 613 SSTYTSPPAPAVSSTYTSPPAPAIISSTYTSPPS-VPVAVSSTYTSPPAPAIIS 670
Qy 1343 TSTPLLPQVPSIPPLQV-ANVPAVOQLHSHQOPALLPNQPTHCEVDS----- 1394
Db 671 TYT-----SSPSVPAVAVSSTYTSPPAPAIISSTYTSPPAPAVSSTYTSPPAPAIIS 726
Qy 1395 ---DTQPKAGIDDKLLEKLSLPSEHSSGQAQHASVLESLVTEIVTQIPTTAYA 1452
Db 727 TYTSSPSAP-----VAVSSTYTSPPAPAIIS-----STYTS-PSAPA 765
Qy 1453 PEKLLTSTSTCL-----PTNLPLGTVALPV-PUVTPQGVSPVSTTSGVKPGAPAK 1507
Db 766 VSTYTSPPSALVLSSTSSPYDIYSPFPAIISGYTSP--SASVAMSSTSSS 822
Qy 1508 PLVTAAPVLPVTELPAGTLPEQLPPPGPSLTOGQPLDLDAQRLRLTLPKXITVTS 1567
Db 823 P-----YDIYVLSASSASRSIATYEFSPSTSLPTS-STYTFSSAYAREFSERSVTS 878
Qy 1568 AVGPVSM-AAPPAITBAGTQPKGVQVEGVAT-----SSGAGVFMKGRQVSV 1618
Db 879 TIAPIQHSTLRLITDFLLQTSMAIQSVSQISTSTLNDIHSALSIVF----- 929
Qy 1619 AADGAQKEGKNKSEDAKSVHFESSSTSESVLSSPESTVPEPGRITPIGSSDVPSS 1678
Db 930 -----NPS-----ASNVTSLIISST-QAISTP-K-NSAKISSLQQLSSS 969
Qy 1679 AHK--TTASEAKSDTQOPTKVGKFOVTTTANKVGRFSVSKTEDKITDTKKGKPVASPPPM 1736
Db 970 TKNPYDTAKNTETGRSTVAVSNFLYTSSAAKPDNEKESATPTEIT----- 1015

QY 1737 DLGQAVLPAVIRPKKEPELSEPSH---LNGSPDPAAFLSRDVGDSGP---HSPHD 1789
 DB 1016 -----TISSSSHAYSLSPSSSHVNTGLSHNFVSSKSATSGYSSSS 1058
 QY 1790 LSSKSLPSQNL--SOSLSN-----SFNSYSSNDSESDIEDLKLRLRLDKLKEI 1841
 DB 1059 ISSIKLSKETIPASKSVSNQERITFTSTLRAN---SQSEKSEGNNSVGSLOSSHSSN 1115
 QY 1842 QDIQSHQKEIESLYTKLGKVPAAV11PPAPLSGRRRRPTKSKSKSRSSSLGNKSPQ 1301
 DB 1116 PSLSTWTKVDSKSLSRKSVSKTNGE-----NGEETGLTTKTQYKSSSETSGSYSRS 1166
 QY 1902 LSGNLGGQASAVLHPOQTLPHPGNTIPESGONLLOPLKSPSSDNLYSAFTSDGAISSP 1361
 DB 1167 FT-KISIGPTTAVQGTASTNSVFTAPALS---TYPTTYP-SPNSYAMLPT--AIIVE 1218
 QY 1962 SLGAPQGSSTNTVATVNSQAAPAMTSS 1995
 DB 1219 S-SETGPTTASFN---PSITGSLPMLEPAVAVS 1248

RESULT 10

S26058
 Probable transforming protein (can) - human

C/Species: Homo sapiens (man)
 C/Date: 13-Jan-1995 #sequence-revision 13-Jan-1995 #text-change 08-Oct-1993

C/Accession: S26058

R/von Lindern, M.; Fornerod, M.; van Baal, S.; Jaegle, M.; de Wit, T.; Buijs, A.; Grosva
 Mol. Cell. Biol. 12, 1687-1697, 1992

A>Title: The translocation (6;9), associated with a specific subtype of acute myeloid le
 k-can mRNA.

A/Reference number: S26058; MUTID:92195315; PMID:1549122

A/Accession: S26058

A>Status: preliminary

A/Molecule type: mRNA

A/Residues: 1-2090 <YON>

A/Cross-references: EMBL:X64228; NID:g29652; PIDN:CAA5535.1; PID:g29653

Query Match 4.4%; Score 475.5; DB 2; Length 2090;

Best local similarity 20.7%; Pred. No. 5.9e-07;

Matches 392; Conservative 238; Mismatches 740; Indels 525; Gaps 79;

QY 545 YVCEGHKTMAKIKDVSLIKRKEQ-----RLVREEDQKKKQESSLSLKQVEQS 596
 DB 309 YIEWDVLVAASASTEVSLIARQSDQINNESWLEDSRAELPYTKSDSLPMGVVD 368
 QY 597 SASQTGI-----KQLPSA-----STGIPTASTTSASVSTQVE--- 628
 DB 369 YTNQVELTIDKTLPPAPVIMLSTDGLCFYMINQNGVMSLITPELSLEGROP 428
 QY 629 -----DEEPPADHQ-----LQYQOPSLISVSDGTVDGQGSVFT---ESR 668
 DB 429 KSPGSTPTTPTSSQAPKLDASAAAAPASLPSSPAPAIATFSLIPAGAPVTFSSSS 488
 QY 669 VSSQGV-----SYGSGHEQAHSTGVPHIPSTVQAGSQPHGVPPSSVA---QGOSQ 719
 DB 489 LKSAATVTGEPSTSSGSD--SKAAPGPEPST-----FSTVPSKSLATTPAASP 538
 QY 720 GQSSSSSLTVGS-----SQPIQHPQOQ--GIQGTAPPO--QTVQVSLSTQTSSEATT 770
 DB 539 VAPAAAFSSFGSSGFKPTLESTPVPVSAPNIAMKSSPSTSAVKVNLSEKPTA--AAT 596
 QY 771 AQVPSQPO--AQVLPQVSAGKOSTOG--VSQVAP--AEVVAQAQOAPTOPTLASVDSAH 826
 DB 597 STPSSSSQAPPMSPFSSASKPAASGSLSHPTPLSAPSSSVPLKSSVLPSPSGSAQSS 656
 QY 827 SDVASGMSDGENPSSSGRHEGRTTKRHVKSVRSRHE--KTSRPLRLTLNNSKGR 885
 DB 657 SPVSMQKSPRTTPAA--KPSPOAKSLQPAVAEKQGHQKWDPM-----AGIGEE 709
 QY 886 VVECOLTHNRKQVTFKFDLDGNDPEELATIMVND-----FILAIER----- 928

DB 710 IAHQKELEBKARTSKACFOVGTSEEMKULRTESDHLFTLEIKETTESLHGDISSLK 769
 QY 929 ---ESF---VDQVREILIEKADEN---LSEDSVVEPEGQGLSELOGKDDYGHSSGOK-- 976
 DB 770 TLLIEGFAGEVEERKEENERNRDSGYLHLYKRLDKSEQLOEIRLHOYKFAVODVN 829
 QY 977 --LEGFKQIPIPASSMPQOIGIP-----TSSLQVWASAGRRF--IYSPVESRLRE 1024
 DB 830 DVLIDEMDQBLEQKKQRHLVPERETLNTLANNHEIINQCKRLNHLVDSIQQLRLYK 889
 QY 1025 -----SKVPSSE-----ITDTVAASTQSPGMNLSHSASSLQOAFSELR 1065
 DB 890 QTSIMELSSAVPQSSIHSDSLSELCNALKTTE-----SHTKSLPKVAKLSPMK 943
 QY 1066 RAQM-----TEBP---NTAPNFSHTGPTFPVPPFLS-----STAGVPTTAAT-- 1107
 DB 944 QAOQLNRLAKRKTPVPRSTAPASLSRSA-----FLSORYEDDERVSTSSVSGSLR 995
 QY 1108 -----APVPATSSPPNDISTSVI-----QSEVTVPTBEGIAG- 1139
 DB 996 SEDARTSCDDEAVQAPRHAPVVRTPS-----IQSLPLHAAAPFAKSHLVHSSPGVMGT 1051
 QY 1140 -VATSTG-VVTSQ-----GLPIP--PV-----SESPVLSV 1166
 DB 1052 SVATSKAKIIPQGADSTMLATKTVKHGAPSPSHISAPQOIAAALRRQWASQAPAVNTL 1111
 QY 1167 VSSI--TIPAVYSISTTSPSLQVP--TSTSEIYVSTALYS-----VTVSATSASAGSTAT 1220
 DB 1112 TESTLKNVPVAVVQELKNNPATPSTAMGSSVYSTAKTTHPVLTTPVANAQAGSLINS 1171
 QY 1221 PGKPPAVVNSQAAGSTVGAATLTSVSTTSPFSTASQSLQSLSSSTPTLAEVTVVSA 1280
 DB 1172 LKSGFTPLPAGQSLSDQKXGCTAKIEAVLSTGASQFSPSPSGFNGFIITPT 1231
 QY 1281 HSLDKTSHSTTGLAFLSLAPSSSSSPGACVSYISQPGELHVLVPSVASTPIILPQA 1340
 DB 1232 PSSNFTAAQAGTSTKSSQPDAPSSGSGSKPEYALPSSPSGGLTASANTTPGEPAA 1291
 QY 1341 GPTSTPLPQVPSIPLVQPVANVPA--VQOTLIHSQOPALLPNQPHTHCEVDSPTOP- 1398
 DB 1292 --SSRPVAPSGTSLSTSSKLETPPSKLGELPPSLAGTLLSFGSLRGQADDSKPT 1349
 QY 1399 -KAPGIDIDKILEKRLSPSEHSSG---AQHASVLETSIVESTVTPGIPPTAVAP 1453
 DB 1350 NKASSTSLSTQPTKSGVSGFNFTAPVLGKHT-----BEPVTSATTSVAP 1399
 QY 1454 SKLLTSTTCLPPTNLPGLTVALPVTVPVTPQVS---TPVST-----TISGVK 1500
 DB 1400 -PATSTSTSTAV-----FG--SLPVTSAGSSGVISFGTSLSAGKTSFSGSQOINTSV 1450
 QY 1501 PGTAPEKPLTKA-----PVLPGTELPAGTLPSEQLPPEPGPS---LTQSQQPLEDL 1550
 DB 1451 PPSAP--PPTTAATPLPFTSPFPLTSFGSLSSATTPS--LMSAGRSITBEATSLAPKPG 1506
 QY 1551 DAQLRRLTSEKXTVTSAVGVMAAFTATTEAGT-----QPOKGVSQVKEGVLTSSG 1605
 DB 1507 DSE-----VSASASLIEEQOQAQLPQAPQOTSDSVKKEPVL----- 1543
 QY 1606 AGVFYKGRFOVSVAAQOQEGKKNKSDASVHPESSTSSSVSSSPSS--TLVKEPN 1664
 DB 1544 -----AQAVNSGTAASSTSLVALSAEATPTTTPVPAPRFAVPPAS 1566
 QY 1665 GITIPG-----ISSDVESAHTKTAASEAKSDT-----GQPKVGRFOVTTANKV-- 1709
 DB 1587 SFSVPGQTAVTAAAISSAGPAVAVETSTPTLASSSTTSIVAQPSAEAAAFGTVSSGVFA 1646
 QY 1710 --GRFVSXTEDKITDTRKKGVPVAPPFMDLEQAVLPAVTPK-----EXPELSEP 1758
 DB 1647 QPAAASSSAFNGQNTNNTATAPASATPVFGVAASTASLRCQCTGSTASTAATAPVSS- 1705
 QY 1759 SHLNGSSDPEAALFLSDVDDGSSPHSPOLSKSLPSQNLQSLNSNFSSTSSGSDNE 1818
 DB 1706 SGFSSPAFGTTAPGVQQTTEGQASV-----FGQSASASAAVSPFSQDPGF 1750

QY 1819 SDIEDEDLKELRRLKDKHKEIQDLSRKHETISLYTKUGKPPAVIIPPAELSGRR 1878
 Db 1751 SSV-----PARGQASS----- 1762
 QY 1879 RRPYKSGS-----KSSRSSSLGNKSPQLSGNLGSGASAVLHPQOTLHPGNTIPES 1930
 Db 1763 -TPSTSSSVGAASSTSSSSSFSGSSPNTGGGLFGQS-----NPAF 1806
 QY 1931 GONULLPLKPSFSSDNLVSAFTSDGATSVPLS-----APGQSTSTNTV----- 1976
 Db 1807 GQ-----SPGQGGGSGVFGGSAITTAATGAFSCASGFGSSNTSGVGOAST 1857
 QY 1977 -GATVNSQAQAQAPPAATSSRKGTFTDHLKLVDMNADANLNGRRSGKHMMVEGPGM 2035
 Db 1858 GGIVFGQSSSSSGSVFSGGNTGRGGFGGGLGKPKSQDANKPFPSSASGCF----- 1910
 QY 2036 ARKSAFAPQLCISMTSNLNGSAPISASATSLGHFTSMCPPOQYGPATPFGAQMGGTG 2095
 Db 1911 -----GSTATSNLTPGN-----GAKTFGGFASSSFGEQR-----PTGTFSSGGG 1952
 QY 2096 GPAPQPLGQFQPVGTASLQNFNISLQKSIISNPPG 2130
 Db 1953 SVASQGFSSSPNKTGFGAARVFGSPPTFGSGPG 1987

RESULT 11

T48814
 hypothetical protein 15B6.220 [imported] - Neurospora crassa
 C:Species: Neurospora crassa
 C:Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 05-May-2000
 C:Accession: T48814
 R/Schulte, U., Alyn, V.; Hohenstein, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura,
 submitted to the Protein Sequence Database, April 2000
 A:Reference number: Z24541
 A:Accession: T48814
 A:Status: Preliminary
 A:Molecule type: DNA
 A:Residues: 1-1952 <SCH>
 A:Cross-references: EMBL:AL353822; GSPDB:GN00112; NCSP:15B6.220
 A:Experimental source: cosmid contig 15B6; strain 74
 C:Genetics:
 A:Gene: NCSP:15B6.220
 A:Map position: 2
 A:Introns: 281/3

Query Match 4.4%; Score 474.5; DB 2; Length 1952;
 Best Local Similarity 20.9%; Pred. No. 5.9e-07;

Matches 446; Conservative 238; Mismatches 748; Indels 699; Gaps 101;

QY 6 AEKQSTPQSLFLSP--PAPAKNGSSSD--SSVGEKLGAAADAADVGTGTEYRRRRHT 60
 Db 166 AADQPTKNGIAHSNPTNRPTDLKGRDTPQVTSV-----ATAKATVG----- 209
 QY 61 MKDSSGAATTTTTHRRFFRSGVTCDSNATALELGLSLPQPSIPAAVPOGAPPEPH 120
 Db 210 -----GAATTTVPGN-----STNSTSFSGGLAMDGTPI-----APFMTQAPP----- 247
 QY 121 REETVATATATQVAAQPPAAAPAGVAGPASTVPSSSTSKORPVSO-----PST 171
 Db 248 -----AVGGMANMDFPST--MDLLOSQSGHILMSTLQSL 280
 QY 172 VSKKEPPPARSGSGGSAKEPOERSQOQDDIEELTRAVGMSNDGRFLKFDIEIGRS 231
 Db 281 SGG-IQPPAAATAATADASLESFALFPAADVFCMTTVALIIGRQR----- 327
 QY 232 FTVVKGADLTETTVAAWCELQDKLTGER-----QRTKEAEMLKGLQHP-----NTV 281
 Db 328 -----AW-----FLAKKEERRAEOYRLKDEYE-AQGILTPAPSEEDAR 365
 QY 282 RFYDSW----- 287
 Db 366 RSKSYIIEEGMLGPESDDEDEGRSPDNRAPKRRKMGVSLPVDISFGAESSLMSA 425

QY 288 ESTVKGKKCIYLVTELM--TSGTLTKYLKRFK-----VMKIVLRSW 327
 Db 426 DQTVDDQKGPVSNQVYSHTPGAAAVNLSALRSPPHHTPTLGIHSPGNIAKATKAI 482
 QY 328 CRQILKGLQFLHT-----RTPEIYHRDKCDNIPTFGTSGYKIGDLGLATLKRAAF 379
 Db 483 SREHLK-IQFNSQAGVEAIFLHNGNFCDEYVHSHDKVLKSGD--RLQVKDVEFVPI 539
 QY 380 ---AKSVYGPPEFNAFMEYEKYDESVDVYAFGCMDEMAISEYSECONAAQIYRRT 436
 Db 540 NGVAEGKTGAEEYEPETPRARYSEGGKMSF-----DFESIHVDNRST 584
 QY 437 S---GVNAPSPDKAIEVEVEIIEGCTRONKDERYSTKDLINLAFQEGTCVRLAEED 493
 Db 585 SPEDNEVMSVD-----PD-RDSDQSELSSEPDDEDILLPDAPDHVM---ET-VEKEBAED 635
 QY 494 DGEKIALKLMRLIEDIK-KLKGKYKDNBEAIEFSDLEBDYVEDVAQEMVESGYCEGDHK 552
 Db 636 DDEAR-----SQSGSVKPEPRERPEPDMSSVPL-----MDPPK-----RQGR 674
 QY 553 TMAKIKDRVSLIKRKEQQLVREGEKKQDESSLKQV-----EQQSASQGTI 603
 Db 675 PPKNGIMSKREELRKKQAMELAKKQPPQPPQEPVYKRGKPRKPRLEPDAADREPR 734
 QY 604 KOLPSASTGIPASTTSASVSTQVREPR-----EADQHQLOLYQOQPSISVLSGT 654
 Db 735 KYPRKKNQEGEADSDAEKTIKERRRKEPTPLELRREDYTEQLQKPNKNVGLIDEV 794
 QY 655 VDSGQS-----SVFSSVSSQQTYSYGS-----QHEQASHSTVPR 691
 Db 795 LSAAPGLLILKQIKYKILQKYPFYFNVDTKGESSVRHLIGNDAFKKEEHLMSRP 854
 QY 692 G-----HDSVQAOQS-----QP-----HGV--YPPSSVAGQS 718
 Db 855 GIDIDAKKKKAPSPHASSLHNFQHYAQPWRPHRGMYHGHVQOQSYHPTVGQRQS 914
 QY 719 ---QGSSSSSLUGVSSQPIQH-----PQQ-----QGIQOTAPRQOQVY 757
 Db 915 YTTTGPG-----ASQHPQHLQTPQPCVPVQOQPPRAYQAATSPRQAQVGTTP 967
 QY 758 -SLQSTSTSEATTAPQ-VSOPAPQVLQVSA-----GKQSTGVSGVAP 801
 Db 968 TAARQMSGTPPATYSSPYVSRPMPFTVAAGSATHSMARQSLPVSQSGQANGIRVAP 1027
 QY 802 AEPVAVAQOQPTTLIASSVDSASDVASGMSDGENVSSSGRHEGRITK----- 853
 Db 1028 -----PPATANTGV-----PVAGGARPAQOATPTANTTAPVRLNPVIAPELIS 1070
 QY 854 -RHVKSIVRSRHEKTSR-PKIRILNYSNNGDRVVECOLETHNRKMTFKFDLDGDP 910
 Db 1071 WLESFVYVEKELEYIKQSKSPQILLAMSVYING-----LKITTKSM-----P 1113
 QY 911 ESIATIMVNDPILAIERESFYDQVRELEKADEMLSDEVSEPEGDQL-----ES 962
 Db 1114 DEESLELV---VLKV---FEERTQGTSHK-----SLDPDLQTLILTFKATMVT 1156
 QY 963 LOGKDDYGGSGQKLEGEFKQPIPASSMPQOIGIPTSSLQVYHSGRRFIVSP---VBE 1019
 Db 1157 LEAKLD-----SQKAE-----CLVLSAIDQVLGLADTKITRGTSEMEREFNAEKVLPA 1206
 QY 1020 SLRESKVPPESEITTVAA---STAQSPGMNLSH---SASSLSLQOAR---SELRRAQM 1069
 Db 1207 ITMKTAEMORKVAAITPAAPVHAITTPGALPANHHTMAPATAPASAAQIATPSATCRANP 1266
 QY 1070 TEG-ENTAPPNFNSHTGPTFPVVPPL---SSIAGVP-----TTAATA-PVPA----- 1112
 Db 1267 PGNFSAVPRITTAAPA-PMPAPLLPRAAPISGHPIASTVNNNMAAGARPVPAVAG 1325
 QY 1113 -----TSSPNDISTSVIOSEVTPPEBGIAG---VATSTGVVTSGLP 1153
 Db 1326 PQIGVGAAGIATPAVRSRPP---VTHGVVTAAPPTTTPRSNLAASVPAQMAQOHTSGAP 1383

QY 1154 -----IPVSHSPVLSVSSITTPAVSISTTSPLOVPTSTSEIVSSNALP 1203
 DB 1384 AGSGNASISRAPSTAPTAFTAPTSVPPVS-----STVRPMSTVPTGPPGLAPPTA--- 1433
 QY 1204 STVSATASAGSSTATPCKPAPVAVSOQAAGSTTGATLTSTVSTTSPSTASQSLQ 1263
 DB 1434 -----SSGASAGTARANAMSTMPAPTSQQAAMTSAPQSVSPRPS---LTTCQGVPA 1485
 QY 1264 SSSTPTPLAETVV-VSAHSLDKTSHSTGTLAFSLAPSSSSPPAG-VSTISQPGI 1321
 DB 1486 IAAASTSRPASGVNPPASSIAPSTK-----SMSEAVPTTASGVNSTVSS----- 1532
 QY 1322 HPLVTSVASTPILPQAAGPTSTPLPQVSPILVQPVANVPAVQQL-THSQPQ- 1377
 DB 1533 -----LAATLPSPA--PRYGPSNNATPTTAPATIPAPLPAAASVSAPVTCQPS 1582
 QY 1378 -PALPNOPTHCPEDVSDTOPKAPGIDIKITLSEKRLSESHSSSGAQAHSVLSL 1436
 DB 1583 APASVPTPP-----PSKLLTSTTSTCLPPTNPLG-----TVA 1476
 QY 1437 VIRSTV--TP--GIPTTAVA-----PSKLLTSTTSTCLPPTNPLG-----TVA 1476
 DB 1611 AISSSIGATPAAISIPSSAPALAPVTVPVQQAASAAARLPVTPAPAAHTIAQSVAP 1670
 QY 1477 LPVT--PVVTPPGVSTPSTTSGVKGTAPSKP-----PLTKAPVLPVTELPAGTLP 1528
 DB 1671 RPTQSEVQSVAGHTSSQASITRAHP-VAOSVPRPNSPTSAAPVAVGTIAQVASAPT 1729
 QY 1529 SEOLPPEPGSLTQSQPLEDLDAQLKRLTSLPEXI-----TWSAVGVSMAAPT 1578
 DB 1730 VTQ--PAPHALSSVSGSL-----POSPHAAQQAAGHTSHASRPVQGVFQ 1775
 QY 1579 AITBAGTQ--PQKGVSGV---KEGPLYATSSGAGVKMGKRFQVSVADQAQEGKRXSE 1632
 DB 1776 SVQATQAVPRBSTALTPTAOPGPVSPAVSGSV-----PAPSAQSVVA--PA 1823
 QY 1633 DAVSHHESSSTSSSVLSSSPSESTLVKPEPNGITIGISSDVESAHKTTASAKSDTG 1692
 DB 1824 PVSTTPVPAATVAPASTVAAAPFTRTVTAAPALSAATNPAPVPSQOHQITGAAPQOQ 1883
 QY 1693 QPTKGRFOVTTTANKVGRFSVSKTEBDKITDTYKKEG--VASPPFMDLEQAVLPVAVIPK 1750
 DB 1884 RPPQAQAAPATPT-----ITSAPPRTPLAPP-----PPPPPPTE 1921
 QY 1751 EKPELSEPSHLNGSPSPPEAFLSRDVGDS 1781
 DB 1922 DPPPPPPPPAEPAPPPPTPLMPTSSAS 1952

RESULT 12

730826
 nascent polypeptide-associated complex alpha chain, muscle splice form gp220 - mouse
 N:Alternate names: alpha-NAC protein
 C:Species: Mus musculus (house mouse)
 C>Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 24-Nov-1999
 C:Accession: T30826
 R:Yocov, M.V., St-Arnaud, R.
 Genes Dev. 10, 1763-1772, 1996
 A>Title: Differential splicing in of a proline-rich exon converts alphaNAC into a muscle
 A:Reference number: Z20889; MUID:96312450; PMID:8698236
 A:Accession: T30826
 A>Status: preliminary, translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-2187 <YOT>
 A:Cross-references: EMBL:U48363; NID:g1666688; PID:g1666689; PIDN:AAH18732.1
 C:Genetics:
 A:Gene: Naca
 A:Map position: 10
 A:Intons: 24/1, 1996/1, 2024/3, 2050/3, 2099/3, 2142/3, 2183/3
 A>Note: differential splicing converts alphaNAC into a tissue-specific DNA-binding activ
 C:Keywords: alternative splicing; DNA binding; transcription factor

Query Match 4.4%; Score 472; DB 2; Length 2187;

Best Local Similarity 21.5%; Pred. No. 7.8e-07;
 Matches 406; Conservative 208; Mismatches 702; Indels 568; Gaps 94;
 QY 533 PEDVAQEMVSGVCE--GDHKTMAKALDRYSLLKRRKQQLVREBEKKQKQSSSLK 590
 DB 340 PEDDKSSAVTNETLSPSSSVNAGTSLSPKASLV-----DKGSNVAIQ 383
 QY 591 QQVEGSSAQQ--TGIKQLPSASTG-----IPTASTSASVSTQVEPEPADQHOQLQYQ 644
 DB 384 PLVTQVPAQKTKGLKEIPSCGATHALDNBSALVMAATVPP----- 429
 QY 645 PSLVSLDGTVDGQGSVFTESRVSQQTVSYSQHEQ-----AHSTGTVP GHIPSTV 698
 DB 430 -----TSSGLVSKDPAASVPTSLVPP-----AAHKQFPAPASATLGVPSPLPATE 476
 QY 699 QASQPHGVYPPSSVAAQGSQGSQSSSILTVSSSQPIQHPOQOQGIQQTAPQQTVOVS 758
 DB 477 GLKNLP-----ISALVNGAPVSPQAQGL-----PTKKDTTLQPLAPALKESSPS 524
 QY 759 LSQTSSEATTAPQVPOAPQVLPQVSAKOSTQGVQVAPAEVAVAAQOATPTTL 818
 DB 525 ASLEVLSEDTYTKTTGGPAPVYRPAT--AGVATTTSLADSP--PAVIRADSCVSPNTV 581
 QY 819 ASSVDSASHDVAGSGMDGENVPSGGRHGRITTKHRYKSVASRHEKTSRPKRLILN 878
 DB 582 SPLKRSVTDPAPAPRTAKNTAPS-----TTSPP--LVP 612
 QY 879 VSNKGRVVECC--LEHNKRMVTFKFDLDGDNBEI-----ATIMVNDPFIATERES 930
 DB 613 LASEGCPVASSNALSPONASVSETALALSPETIKSVFPDPPLAELSFST-----AR 664
 QY 931 FVDQVEIIEKADEMLSE--DVSVEPEG-----DOGLE--SLQKDDYGFSGQKLEGEFK 982
 DB 665 KUDAVHMESSGSRQGHDPASVTAKGTVCCLADSLDTSVSAKSGSALSGAASP-----L 720
 QY 983 QPIPASMPQ--QIGIPYSSLTQVHSAGRITVSPVPSRLRESGVPSSETTDYVAA 1039
 DB 721 YPLEVFLPAGLAQVQPKSLNK-----LSPTPS--SKGAP--VPS 759
 QY 1040 TASP-----GNNLSHASSLSLQAASEL-----RAQMEGNTAPAPNPSHT 1083
 DB 760 TGAPPEFKGAPVPTBESSISKQVPAELPSQKTEPVTASRLISAVQSK--VDFIMSDV 818
 QY 1084 GPTFP-----VPPFLSSIAGVPTTAAT--APVATSSPPNDISTVTSQSEVTVTEE 1135
 DB 819 TPTSPKTSATAPKOTSATLSLKSVAVATSLSPKAPVAPSE--ATVPTETLPTSLKN 876
 QY 1136 GIAGVATSTGVVTSGLPIPPVSESPVLSVSSITL--PAVSTSTTS----- 1182
 DB 877 ALAAATPKETLATSS--IPKVT--SPSPQKTPKSVSLKGAPAMTSKKAETIAASKDVSPS 931
 QY 1183 -----PSLQ--VPTSTSEIVSST--ALY-----PSVTSAT-----SASAG 1216
 DB 932 QPPEVPLLQHPPTSPKSPVSDTLGALLTSPPKPPATIAETPTTPPKSPKPAASK 991
 QY 1217 STATPPGK-----PP-----AVVSOQAAGSTVGATLTSTVST-----TTSFP 1253
 DB 992 TPATPPEGTAVPLEIPCKSKAPKTAAPKESATSSKRAPXTAVSKETPSKGVAVP 1051
 QY 1254 STASQSLTQSSSTSTYTLAETVVVSAHSLDKTSHSTT--GLAF----- 1286
 DB 1052 ---LEIPLPKETISKSAIPGKKSASSPKRSPTKAGTPEPGGVAVPPEISLPPKETPQ 1108
 QY 1297 -----SLSAPSSSSSGAGVSVYSISQPGG--HPLVTPS----- 1328
 DB 1109 NATPNESIAASSQGRSPKTSVYPKK--TPPGGVTAAPLEIPAPAPQAPATVAPKQIPTEDA 1167
 QY 1329 --VIATPILPQAAGPTSTPL--LPQVPSIPLVQPVANVPAVQQLTHSQPQ--PALIP 1382
 DB 1168 VTIIAGSLSPKSKAKTAAPKAPATPSGVIVAVSGEISPEPKTKSTAAAPKENSATLPP 1227
 QY 1383 -NORPHCPEDVSDTOPKAPGIDIKITLSEKRLSESHSSSGAQAHSVLSLSTSL----- 1436

Db 549 -----ASTTAGSALAASA-----ANAKHQ-----SSAPSSGSSS- 579
 QY 663 VETESKSSQOTVSVYSGHQHAGHSTGTVEGHLPTVQAOSQPHGVYPPSSVAQSGSQCP 722
 Db 580 -----SRRSSQN-----DAATAAG-----GIYVMSGTRHGGV-----QNRAP 613
 QY 723 SSSSLTGVSSSOPICHPQOQGGIQQTAAPQQTQVYLSQSTSSSEKTTAQAQVQPAQV 782
 Db 614 TSKQAT-ILSLDP-----PSYKPSNTTQIAQIPLEFNKSTATSSAA----- 655
 QY 783 LPOVSAKOSTQGVSAQVAPAEVVAQPCATQPTTLASVDASHSDVASCMDGNGENVPS 842
 Db 656 ---QPSIGICTR-----KIADPKGRIP--LNTAVQGHRTATGAVVAANGGIBS 700
 QY 843 SSGRHEGRTTKRYKYSVRSRSHKTSRPKRIILVNSKNGRVNVECOLETHNRKMYTK 902
 Db 701 ---HRDHAQOQOQYMQQLTSTTMMSK-----LINKTPAAGGTAAATSSSSSSATSTAP 749
 QY 903 FDLIDGNPEELIATIMNNDITLAIERESFVDQVREIIEKADMLSEDPVSEPPGDOGLS 962
 Db 750 LQKSGQISHAPTEPV-----IRE-----DDENNSKQNGNVP 783
 QY 963 LQKXDYGFSGSQKLEGEFKQPIPASSMPQOIGIPTSSLTQVHVSAGRPVSPVESRL 1022
 Db 784 LIG-----GVGPQTSFPAVQVPTEDATSSSDKEQOQKAS-----SETP----- 821
 QY 1023 RESKVPSEITDVAASTAOSPGKMLSHSASSLSLQAFSELRAQMTREGPNTAPPFESH 1082
 Db 822 KSKK--PSMT-----HOSPMPSPSQMTAMEBLKLSSEGGQGTGPVAT 862
 QY 1083 TGTTFPVVPPFLSLAGVPTT--AAATAPVATSSPNDSITSVIQLSEVTPTEEGIAGV 1140
 Db 863 GGGPQAPATSQOMBSATTTNANAGASSGGAAMAAATNQLS-----GA 905
 QY 1141 ATSTGVATSGGLPIPVYSESPVLSSVVSITTPAVSISTTSPSLQVPTSTSEIVASTA 1200
 Db 906 PPSGTASSQGYHPRAPSSSSSSSTNPHQQLTNHASFVSPTSSYQIPTSTA-VNVTSTG 964
 QY 1201 LYPSVTYSA-----TSASAGSTATP--GKPRPAVSSQAAAGSTTVGATLT 1244
 Db 965 MPTSSSSSAPPRTNRKOTFHGKTEKXGDDSDDEIGETP-----GNVSIATGP 1015
 QY 1245 SVSTTSPFPTASQLSIQLSSSTPTLAEFTVVAVSASHLDTSHSSTGLAFLSAPSSS 1304
 Db 1016 SANAADA--TWSKLSKLTTRDNRRESMTQFVSGRAGTIGASQOQTRAAALAIRESQSP 1073
 QY 1305 SSPGAG 1310
 Db 1074 IAPGAG 1079
 RESULT 14
 T34513
 Hypothetical protein ZK783.1 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
 C:Accession: T34513
 R:Favelli, A.; Vaudin, M.
 Submitted to the EMBL Data Library, August 1994
 A:Description: The sequence of C. elegans cosmid ZK783.
 A:Reference number: Z21536
 A:Accession: T34513
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-1507 <FAV>
 A:Cross-references: EMBL:U13646; PDB:AAC24418.1; GSPDB:GNO0021; CESP:ZK783.1
 A:Experimental source: strain Bristol N2; clone ZK783
 C:Genetics:
 A:Gene: CESP:ZK783.1
 A:Map position: 3
 A:Introns: 14/1; 48/2; 84/1; 196/3; 303/1; 381/1; 586/1; 605/1; 1175/3; 1207/1; 1409/2;
 3504/1

Query Match 4.3%; Score 467.5; DB 2; Length 3507;
 Best Local Similarity 19.9%; Pred. No. 1.7e-06;
 Matches 425; Conservative 278; Mismatches 706; Indels 723; Gaps 100;
 QY 1 MSGGAEEKOSTPGSIFLSPAPAPKNGSSSDSVCEKIGAAAADAVTGTEYRRRRHT 60
 Db 1224 VDGDNGETSGVDG---PPTPAPPPSSASBSTSRIPPTSSASDESGGE-----G 1273
 QY 61 MDKDSGAAATTTTTHFRFRSVICDSNATLALBGLPLSLPQSPIPAAPQAPPEPH 120
 Db 1274 VPSPDGSSESTSPAD-----GVSPSSATAPEVTTTAS-----STPDVSSGIP--- 1321
 QY 121 REETVATATSGVAAQPPAAAPGEGAV---AGAPSTVP-----SSTSKDRPVYQPS 170
 Db 1322 ---STSKTAPLPLETTAPSTEVTPREGSGTEESTLPPEGSGESTTSGAPVTEPA 1373
 QY 171 LV---GSKEPPPAR-----SSGGGSAKPPQERSQ---QOD----- 202
 Db 1374 TVLPQNRNEKEPPTKDTPTALPTTTGAPQANDSVENKTKTSSDRCGLALCBERTGVC 1433
 QY 203 -----DIELETKAVGMSNDR-----FLKPDIEIGRS 231
 Db 1434 KCPGFEAGAPPKKSCVDVDECATGDHNCHEASACQVYVGACFCPTGFFKAD---DGS 1489
 QY 232 FKTIVKGLDETTEVA---WCELDQRL-----TKSEROR 264
 Db 1490 CQDIDECTENHSTCCGANAKCVNKPQTVSCENGFAGDQCVPTTKPKPCDSTOSKSH 1549
 QY 265 FKEAEKMLKQLQHPNIVRFDSWESYVKGKCIIVLTMLTSLTKLYLRFYMK----- 320
 Db 1550 CSBS-----NMSCEVDIVDSVECKEC-----WGYKKSAGVCEIDNE 1587
 QY 321 -----IKLIRSM---CRQLKLOPLMT-----RPPILHRDLKDNIF 356
 Db 1588 CVAEKAPCSINAMCVNNGTFPSCCKQGYGDDGMCTDINECERHP--CHPAECTNL- 1644
 QY 357 ITGPTGSKVI-----GDIGL-----ATLRASAKSVIG 385
 Db 1645 ---EGSFKECHSGFEGD--GIKKTNPILERSCEDVEKFCGRVDHVCLSVRIINGSLSS 1699
 QY 386 TPPEMAPEMTEBKYDESDVYVAFGMCMLEMAISEPYSECONAQIYR-RVTSGVKPAF 444
 Db 1700 VCECEPGEFREKESNSCVDI--DEGESRNCPDASAVCVNTGTGSTRCEAGYBEGG 1756
 QY 445 DKVAIPVKEIIEGCIQRNDEYSIKDLNH-----AFQEBTGVRLAEEDDG 495
 Db 1757 VCTDIDECDEKMGAC-----DSYAMCINRMSCCGCKMAGYTGAGATCIKIEBPKS 1808
 QY 496 EKIA-IKLMRLIEDIKK-----LKRYKDN--EALIEFS----- 525
 Db 1809 DKTAGCTDEMSRLCELEKKQCTVDEEVPQCGACLPGRHPIINGTCQSLQISGLCAQKNDON 1868
 QY 526 -----FLDERD-----VPE-----DVAQEVESGVYCEBPHKTMARIK--DRVSL 564
 Db 1869 KHAECIDHPSHFCSCPDGFTGDMICDVBECNNAQ--WDDENTKCENTIGSFNCVCL 1927
 QY 565 IKRKEOROLVREBEQKKESSSLKQVBOSSAQIGIKQLPSASTGIPASTTASASV 624
 Db 1928 EGFKKVDEKCVADKKQPNRKEIIDE--EMSSSSNG--QEKPTTKGIVSTASTSSES 1983
 QY 625 TQVEPE-----EPADQHQQLQYQPPIS-----VLSDTGVNSGQ-- 660
 Db 1984 TTAEPHYTTSISTTSTKMTSSKSPENVTWSSSPVSTSSSKSTTASSETTVSTPSES 2043
 QY 661 -----SSVFTESRYSQOTVSVYSGHQHAGHSTGTV-----PGHPSVYQAOSOPHG 706
 Db 2044 SSSSEAPLTSAPATTEVTESVSKSTTKKSSSEITVYKLSKSBVEVSSV--KSPTS- 2100
 QY 707 VYPPSVYAQOSQOQPSSSSLTGVSSSQPI--QHPQOQGGIQQTAAPQQTQVYLSQSTST 764
 Db 2101 --TSTTISQSVTSYVPESTKSTVLSSEAPVATSTPTEVHTSSETKP-----SISASST 2151
 QY 765 SSEATTAP-----VSQPAQOVLPQVSAKQSTQGV--QVAPAEVVAQAQAPATPT 816

```

Db      2152 TGDNTSTPTSTSLASVSTAP-----EGTASAVAVK-LSSLSPVQSPS 2197
QY      817 TLASVDSAHSDVASGMSDGNENVPSSGRHGRITTKRHYKSVRSRHEKTSRPKRI 876
Db      2198 TKTFPATESSTVQASSETSSGT-----SVKS-----TSEP----- 2226
QY      877 LNVSKKGRVVECCOLEFNKRKMYTFKPLDGDNPBEIATIMVNDPILAIERESFVDQR 936
Db      2227 -----ESH-----VT-KLSITSNPS--SSVPTSP----- 2249
QY      937 EIEKADEMLSEDEVVEPEGOGLESLOKXDYGSGSQKLEGERKQIPASSMPQOLCI 996
Db      2250 KSTPVPSTEQPTSTPSG-QSLTPMN-----SNSEVLITSEPHYLSSLSLD--V 2238
QY      997 PTSSLTQVHSGRRFIVSPVESRLRESKVPSEIT-----DTVAASTASQPMNLH 1050
Db      2299 SOSSTTP-----NNLSSTVETPTKTSSEVSLNSRPTTEPTLSPDI-LST 2346
QY      1051 SASLSLQOAFSELRRAOM-----TEGNTAP--PNESTGPTFPVPV----- 1091
Db      2347 TTNMLSSQSTVSTDRSEISENSEKPTSADELVTASVTHVASSPDPVPTSSRPDDLGT 2406
QY      1092 -----FPLSS--IAGVP--TTAAATAPVPATSSPNDIST--SVTQSEVTVPTBGI 1137
Db      2407 SSTNITPBAASKOTISTPTPTDTTASBEPKYSTWSP-DLSTTSNVLSESTPSESSK 2465
QY      1138 AGVATST--GVVTSGGI-----PIP-----PVSESP 1161
Db      2466 SPVASSSTEGISVVTSTEFKVPSTISSVLEEDLTKTTPSTILETTASTSTSPLEDS 2525
QY      1162 VLSSV-----VSSITTPAVAVSISTSPSLQVPTSTSEIVSSTALYPSVATASAS-- 1213
Db      2526 LTVSVRIHELTTSENPKSESESTTSSSESKPQEPAGILTSVVVPTSVTSILTASEI 2585
QY      1214 -----AGGSTATPGPK-----PPAVSQQAAGSTVGAALTSVSTTSPST 1255
Db      2586 EATISNTPFOKQRTPTTSKSLVKTSTSPSTVSSPSSTKRTTVSTVSTTPTHEET 2645
QY      1256 ASQSLISQSSSTPTPLAETVVAHSLDKTSHSTTGLAFSLAPSSSSSPG---AGV 1311
Db      2646 TTSSELTTLTAPSKPT-----ESTTESSEPTTPATSKETKSNVSTSRKSTENV 2696
QY      1312 SSTISQPGHPLVIVSVIASPILPQAGPTSTPLIPQVPSLPPVQVAVNPAVOQLT 1371
Db      2697 ETSTSSQSGSL-----SSTWSTSS-----SEPTNAPAV--T 2726
QY      1372 IHSQPOFALLPQPHHCPVVDSDTQPKAPGIDDKLLEKASLSEHSSGA----- 1425
Db      2727 VSSEASTTLEENSSSTSP-----TSSEA-----SVKLSLTPRESITSEAVTVSSR 2772
QY      1426 -----QHASVSL-----ETSLIVESTVTPGIPPTVAVAPSK--LITSTSTCLP 1467
Db      2773 APAEITWSSSHREISTVSESEPEPELPLSTVSPNVATASLSPSEPIILSTSSSTPR 2832
QY      1468 TNPPLGT-----VALP-----VTPVVTGQVST--PVSTTSGV 1499
Db      2833 VRLITGTPDDLIVSVVPSHGNRRONITASSVPSNSTSPILIPSESLTTPQPEPTTTTA 2892
QY      1500 KPCTAASK--PPLTKAPVLPGTELPAGTLPSEQLPPFQPSLTSQOQPELIDAOILRT 1557
Db      2893 KPAITSGKGFPSIQPAEMFTTPAP--PPSNGVGEETNOEE-----CVTST 2940
QY      1558 LSEKXITVTSVAVPVSMAAPALTEAGTOPKGVSVQKGPVLATSSGAVFMGRFQV- 1616
Db      2941 TTTEAPSLCSTV-----TCHSLATCOSTGVCICRGFT-----GGTTACSKSKTA 2987
QY      1617 -----SVAADQAKGKNSKSEDAKSVHESSTSESVLSSSPSTLVKP--EPNGI- 1666
Db      2988 DCTSLPDLADRAKCNSTRSCGC-----DAGYIGGVYCSPPHQCVCVRLDNLCSPEANC 3042
QY      1667 -----TIPGSSDVE--SAHKTTASEAKSD 1690

```

```

Db      3043 QNRRCQCLPGFTGDGVKCVSIHERASNCQCD 3074
RESULT 15
548478
glucan 1,4-alpha-glucosidase (EC 3.2.1.3) - Yeast (Saccharomyces cerevisiae)
N/Alternate names: extracellular glucocanylase; mucin-like protein MUC1; protein YIR019C
C/Species: Saccharomyces cerevisiae
C/Date: 10-Sep-1999 #sequence revision 10-Sep-1999 #text_change 12-Nov-1999
R/Accession: 548478; A26877; E26877; S27281; J06123
R/Romley, K.
submitted to the EMBL Data Library, October 1994
A/Reference number: 548478
A/Accession: 548478
A/Molecule type: DNA
A/Residues: 1-1367 <ROW>
A/Cross-references: GB:247047; EMBL:238061; NID:9603997; PID:9763364; GSPDB:GNO0009; MIM:
R/Yamashita, I.; Nakamura, M.; Fukui, S.
J. Bacteriol. 169, 2142-2149, 1987
A/Title: Gene fusion is a possible mechanism underlying the evolution of STA1.
A/Reference number: A91831; MIM:87194600; PMID:3106330
A/Accession: A26877
A/Molecule type: DNA
A/Residues: 1-242 <YAM>
A/Cross-references: EMBL:M16164; NID:9172522; PIDN:AAA35014.1; PID:9172525
A/Accession: B26877
A/Molecule type: DNA
A/Residues: 762-1331 <YAM>
A/Cross-references: EMBL:M16165; NID:9172523; PIDN:AAA35015.1; PID:9172526
R/Pardo, J.M.; Ianez, E.; Zalacain, M.; Claros, M.G.; Jimenez, A.
FEBS Lett. 239, 179-184, 1988
A/Title: Similar short elements in the 5' regions of the STA2 and SGA genes from Sacchar
A/Reference number: S27281; MIM:89031230; PMID:3141213
A/Accession: S27281
A/Molecule type: DNA
A/Residues: 1-31 <PAR>
A/Cross-references: EMBL:X13857; NID:94551; PIDN:CAA32069.1; PID:94552
R/Lambrecht, M.G.; Bauer, F.F.; Marmur, J.; Pretorius, I.S.
Proc. Natl. Acad. Sci. U.S.A. 93, 8419-8424, 1996
A/Title: Muc1, a mucin-like protein that is regulated by Mss10, is critical for pseudohy
A/Reference number: J06123; MIM:96323237; PMID:8710886
A/Accession: J06123
A/Status: nucleic acid sequence not shown
A/Molecule type: DNA
A/Residues: 1-1167 <LAM>
A/Cross-references: GB:U03062; NID:91304386; PIDN:AA49609.1; PID:91304387
C/Genetics:
A/Gene: SGD:MUC1; STA2; MAL5; DEX2; SGD:S0001458
A/Cross-references: MIPS:YIR019C; SGD:S0001458
A/Map position: 9R
C/Superfamily: yeast glucan 1,4-alpha-glucosidase homolog; glucan 1,4-alpha-glucosidase
C/Keywords: glycosylase; hydrolase; polysaccharide degradation; transmembrane protein
F:350-1367/Domain: transmembrane #status predicted <TM1>
F:350-1367/Domain: transmembrane #status predicted <TM2>
Query Match 4.3%; Score 460; DB 1; Length 1367;
Best local Similarity 22.4%; Pred. No. 1,1e-06;
Matches 302; Conservative 164; Mismatches 540; Indels 342; Gaps 54;
QY      585 EESSLKQVQVQSS--ASQGTIKOLPSASTGIPATSTAS-----VSTQVEPEPEAD 635
Db      229 EESTTSTSTSESTSTSTSTSTSTSTSTTAATPTTSTCKKRPPTTSTCKEPTTP 288
QY      636 QHQOQLOQOPSLISVSDGTVGDSQGSQSVTERVSSQOQTVSGSQHEQAHSTGYVGHIP 695
Db      289 HND-----TTPCTKKKTTTSTKCT-----KTTTTPVPTPS 319
QY      696 STVOAQSOPHGVYPPSSVVAQSQSQSQSQSSSILTVSSSQPIQHPOQSQSQIQTAPPOQTV 755
Db      320 STTBSSAP--VTPSSSTTSSSAPVTSSTT-BSSAPVPTPSSSTTBSSAP----- 370
QY      756 QYSLQSTSTSEATTAQPVSPQAPQVLPVVSAGKQSTQGVSQ--VAPAEVAVAAQQA 812

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 24, 2004, 01:11:09 ; Search time 204 Seconds

(without alignments)
3366.909 Million cell updates/sec

Title: US-10-010-720-14

Perfect score: 10812
Sequence: 1 MSGAAKQSTPGSLFSP.....NLSIKQISNPSGSLRTT 2136

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1349238 seqs, 321558718 residues

Total number of hits satisfying chosen parameters: 1349238

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

```
1: /pubshed/Applications/AA:*
2: /cgn2_6/ptodata/2/pubppaa/US07_PUBCOMB.pep:*
3: /cgn2_6/ptodata/2/pubppaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubppaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubppaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubppaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubppaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubppaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubppaa/US09_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubppaa/US09C_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubppaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubppaa/US10_NEW_PUB.pep:*
13: /cgn2_6/ptodata/2/pubppaa/US10_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubppaa/US10_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubppaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubppaa/US10C_NEW_PUB.pep:*
17: /cgn2_6/ptodata/2/pubppaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/2/pubppaa/US60_PUBCOMB.pep:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	10794.5	99.8	2135	US-10-362-892-9	Sequence 9, Appli
2	10794.5	99.8	2135	US-10-288-798-9	Sequence 9, Appli
3	10677	98.8	2382	US-10-336-472-230	Sequence 230, App
4	10677	98.8	2382	US-10-196-935A-2	Sequence 2, Appli
5	10677	98.8	2382	US-10-052-648A-40	Sequence 40, Appli
6	10677	98.8	2382	US-10-052-648A-1404	Sequence 1404, Ap
7	9204	85.1	2126	US-10-052-648A-39	Sequence 39, Appli
8	3408	31.5	670	US-10-052-648A-38	Sequence 38, Appli
9	2439	22.6	1743	US-10-460-545-2	Sequence 2, Appli
10	2130.5	19.7	1345	US-10-433-794-17	Sequence 17, Appli
11	2126	19.7	1231	US-10-052-648A-37	Sequence 37, Appli
12	2126	19.7	1231	US-10-196-935A-4	Sequence 4, Appli
13	2072	19.2	453	US-09-925-301-1286	Sequence 1286, Ap
14	2042.5	18.9	1251	US-10-114-270-80	Sequence 80, Appli
15	2038.5	18.9	1084	US-10-415-011-17	Sequence 17, Appli

16	2038.5	18.9	1234	US-10-052-648A-36	Sequence 36, Appli
17	1959.5	18.1	560	US-10-114-693-84	Sequence 84, Appli
18	1866.5	17.3	1356	US-10-114-270-78	Sequence 78, Appli
19	1861	17.2	359	US-10-654-421-152	Sequence 152, App
20	1804.5	16.7	779	US-10-353-929-49	Sequence 49, Appli
21	1776	16.4	1069	US-10-182-243-50	Sequence 50, Appli
22	1677	15.5	324	US-09-933-767-533	Sequence 533, App
23	1677	15.5	324	US-10-004-860-533	Sequence 533, App
24	1677	15.5	324	US-10-023-282-533	Sequence 533, App
25	1677	15.5	324	US-10-023-282-533	Sequence 533, App
26	1572.5	14.5	557	US-09-862-027-40	Sequence 40, Appli
27	1362.5	12.6	663	US-10-052-648A-12	Sequence 12, Appli
28	1323	12.2	309	US-10-108-260A-3492	Sequence 3492, Ap
29	1280.5	11.8	1381	US-10-333-131-25	Sequence 37, Appli
30	1280.5	11.8	1381	US-10-333-131-25	Sequence 37, Appli
31	761.5	7.0	613	US-10-240-145-77	Sequence 25, Appli
32	745.5	6.9	516	US-09-862-027-39	Sequence 39, Appli
33	745.5	6.9	516	US-09-862-027-39	Sequence 39, Appli
34	736	6.8	587	US-10-437-963-175930	Sequence 34, Appli
35	729	6.7	567	US-10-424-599-222750	Sequence 222750, A
36	723.5	6.7	618	US-10-425-114-69051	Sequence 6051, A
37	722.5	6.7	746	US-10-437-963-164986	Sequence 164986, A
38	706	6.5	604	US-10-425-114-72414	Sequence 72414, A
39	702.5	6.5	549	US-10-424-599-261688	Sequence 261688, A
40	694.5	6.4	477	US-10-425-114-64415	Sequence 64415, A
41	688	6.4	417	US-10-437-963-144326	Sequence 144326, A
42	683	6.3	658	US-10-424-599-170860	Sequence 170860, A
43	681	6.3	864	US-10-437-963-183335	Sequence 183335, A
44	678.5	6.3	677	US-09-862-027-38	Sequence 38, Appli
45	678.5	6.3	678	US-10-437-963-183341	Sequence 183341, A

ALIGNMENTS

```
RESULT 1
US-10-362-892-9
; Sequence 9, Application US/10362892
; Publication No. US20040038861A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.; BANDMAN, Olga
; APPLICANT: NGUYEN, Daniel B.; WALIA, Narinder K.
; APPLICANT: HAFALIA, April J.A.; YAO, Monique G.
; APPLICANT: GANDHI, Ameena R.; GURURAJAN, Rajagopal
; APPLICANT: DING, Li; PATTERSON, Chandra S.
; APPLICANT: YEE, Henry; BAUGHN, Mariah R.
; APPLICANT: TRIBOULEY, Catherine M.; THORNTON, Michael B.
; APPLICANT: ELLIOTT, Vicki S.; LU, Yan
; APPLICANT: ISON, Craig H.; AU-YOUNG, Janice K.
; APPLICANT: TANG, Y. Tom; AZIMZAI, Valda
; APPLICANT: BURRILL, John D.; MARCUS, Gregory A.
; APPLICANT: ZINGLER, Kurt A.; LU, Dying Aina W.
; APPLICANT: LAU, Preeti G.; RAMKOWAR, Jayalakshmi
; APPLICANT: WARREN, Bridget A.; KEARNEY, Liam
; APPLICANT: POLICKY, Jennifer L.; THANGAVELU, Kavitha
; APPLICANT: BURFORD, Neil
; TITLE OF INVENTION: HUMAN KINASES
; FILE REFERENCE: PF-0209 USN
; CURRENT APPLICATION NUMBER: US/10362, 892
; CURRENT FILING DATE: 2003-02-25
; PRIOR APPLICATION NUMBER: PCT/US01/27219
; PRIOR FILING DATE: 2001-08-31
; PRIOR APPLICATION NUMBER: US 60/229, 873
; PRIOR FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: US 60/231, 357
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: US 60/232, 654
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: US 60/234, 902
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: US 60/236, 499
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: US 60/238, 389
```

PRIOR FILING DATE: 2000-10-06
PRIOR APPLICATION NUMBER: US 60/240,542
PRIOR FILING DATE: 2000-10-13
NUMBER OF SEQ ID NOS: 48
SOFTWARE: PERL Program
SEQ ID NO 9
LENGTH: 2135
TYPE: PR
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: Incyte ID No. US20040038881A1 5502218CD1
US-10-362-892-9

Query Match 99.8%; Score 10794.5; DB 12; Length 2135;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2134; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

1 MSGGAARKOSTGSLFLSPAPAPKNGSSSDSVGEKLGAAAADAVTGRTEYRRRRHT 60
1 MSGGAARKOSTGSLFLSPAPAPKNGSSSDSVGEKLGAAAADAVTGRTEYRRRRHT 60
61 MDKDSRGAATTTTTEHFRFRSVICDSNATALELPLPLSLPQPSIPAAVPOGAPPEPH 120
61 MDKDSRGAATTTTTEHFRFRSVICDSNATALELPLPLSLPQPSIPAAVPOGAPPEPH 120
121 REETVATATSVQAQOPPAAPAEQAVAGPAPSTVSSSTSKDRPVQPSIVSGKEPPP 180
121 REETVATATSVQAQOPPAAPAEQAVAGPAPSTVSSSTSKDRPVQPSIVSGKEPPP 180
181 ARSSGGGSAKEPEERSQOQDDIEELETKAVGMSNGRFLKPIELIGRGSFKTVYGLD 240
181 ARSSGGGSAKEPEERSQOQDDIEELETKAVGMSNGRFLKPIELIGRGSFKTVYGLD 240
181 ARSSGGGSAKEPEERSQOQDDIEELETKAVGMSNGRFLKPIELIGRGSFKTVYGLD 240
241 TETVEVAMCELOQRKLTSEKRFKEAEMLKGLQHPNIVRFYDSVESTVXGKCIIVLV 300
241 TETVEVAMCELOQRKLTSEKRFKEAEMLKGLQHPNIVRFYDSVESTVXGKCIIVLV 300
241 TETVEVAMCELOQRKLTSEKRFKEAEMLKGLQHPNIVRFYDSVESTVXGKCIIVLV 300
301 TELMTSGTLTKYKRFKVMKIVLRSCQQLKGLQFLHTTPPIIHRDLKCDNIFFITGP 360
301 TELMTSGTLTKYKRFKVMKIVLRSCQQLKGLQFLHTTPPIIHRDLKCDNIFFITGP 360
301 TELMTSGTLTKYKRFKVMKIVLRSCQQLKGLQFLHTTPPIIHRDLKCDNIFFITGP 360
361 TGSVKIGDLGATLKRAFAKSVIGTPEFMAPEMTEEKIDSVYAFAGCMLEMAISEY 420
361 TGSVKIGDLGATLKRAFAKSVIGTPEFMAPEMTEEKIDSVYAFAGCMLEMAISEY 420
421 PYSECQNAAQYRRVTSGVKPASFDKVAIPVKEIIIEGCIKONKDERYSIXDLNHAFFQ 480
421 PYSECQNAAQYRRVTSGVKPASFDKVAIPVKEIIIEGCIKONKDERYSIXDLNHAFFQ 480
481 BETGVAVELAEEDGEEKIAIKLMLRIEDIKLKGKYKDNEMAIFFSFDLERVPEEDVAGEM 540
481 BETGVAVELAEEDGEEKIAIKLMLRIEDIKLKGKYKDNEMAIFFSFDLERVPEEDVAGEM 540
541 VESGVYCEGDHKTAKAIKDRVSLIKRREOROLVREOEKKOEKKESSLKOOVESSASQ 600
541 VESGVYCEGDHKTAKAIKDRVSLIKRREOROLVREOEKKOEKKESSLKOOVESSASQ 600
541 VESGVYCEGDHKTAKAIKDRVSLIKRREOROLVREOEKKOEKKESSLKOOVESSASQ 600
601 TGIQOLPASTGIPASTTSASVSTOVEPEEPEADQHQOLQYQPSISVLSGTVDSGQ 660
601 TGIQOLPASTGIPASTTSASVSTOVEPEEPEADQHQOLQYQPSISVLSGTVDSGQ 660
661 SSVFTESRVSQOQVSYSGOHEQAHSTGTPGHIPSTVOAQOSQPRGAVFPSSVAQSGQ 720
661 SSVFTESRVSQOQVSYSGOHEQAHSTGTPGHIPSTVOAQOSQPRGAVFPSSVAQSGQ 720
721 QPSSSITGVSSSQOIQHPOQOQOIQOTAPQOQVYSLSOTSTSEATTAQOPVQOPAP 780
721 QPSSSITGVSSSQOIQHPOQOQOIQOTAPQOQVYSLSOTSTSEATTAQOPVQOPAP 780
721 QPSSSITGVSSSQOIQHPOQOQOIQOTAPQOQVYSLSOTSTSEATTAQOPVQOPAP 780
780 QVLPVSAKQSTQGVSVQAPAPAVQAPATQPTTLASSVDASHDVAASMSDGENV 839

841 PSSGRHEGRTTKRHRKSVRSRREKTSRPLRLINVSNGDRVVECOLETHNKKMT 900
840 PSSGRHEGRTTKRHRKSVRSRREKTSRPLRLINVSNGDRVVECOLETHNKKMT 900
901 FKEDLDGNDPEEJATIMVNDPILAIERESFVDOVEIIEKADENTSEDSVEBQDGL 960
900 FKEDLDGNDPEEJATIMVNDPILAIERESFVDOVEIIEKADENTSEDSVEBQDGL 960
961 ESTQKDDYGFSSQULBEGFKQPIPASMPQOIGIPSSLTQVHSAGRRFVSPVES 1020
960 ESTQKDDYGFSSQULBEGFKQPIPASMPQOIGIPSSLTQVHSAGRRFVSPVES 1020
1021 RLRESKVPSEITDVTAASTAGSPGMNLSSASLSLQAFSELRRQOMTEGNTAPPNF 1080
1020 RLRESKVPSEITDVTAASTAGSPGMNLSSASLSLQAFSELRRQOMTEGNTAPPNF 1080
1081 SHTGPTFPVVPPLSSIAQVPTTAATAPVPATSSPPNDISTVYQSEVTVPEBGIAGV 1140
1080 SHTGPTFPVVPPLSSIAQVPTTAATAPVPATSSPPNDISTVYQSEVTVPEBGIAGV 1140
1141 ATGTGVTSGGLPIPPVSESPLVSSVSTITPAVVISITSPSLQVPTSETVSSSTA 1200
1140 ATGTGVTSGGLPIPPVSESPLVSSVSTITPAVVISITSPSLQVPTSETVSSSTA 1200
1201 LYSVTVSATSAAGSGTATPGKPPAVVSOQAAGSTTVGATLTSVSTTSPBSTAQLS 1260
1200 LYSVTVSATSAAGSGTATPGKPPAVVSOQAAGSTTVGATLTSVSTTSPBSTAQLS 1260
1261 IOLSSSTPTLAEVTVVSAHSLDKTSHSSSTGLAFSLAPSSSSSGAGVSSITSPGG 1320
1260 IOLSSSTPTLAEVTVVSAHSLDKTSHSSSTGLAFSLAPSSSSSGAGVSSITSPGG 1320
1321 LHPLVIFSVASAPPILOAAGPTSTPLPOVPSIPPLVOPVANYPAVQOILHESQPPAL 1380
1320 LHPLVIFSVASAPPILOAAGPTSTPLPOVPSIPPLVOPVANYPAVQOILHESQPPAL 1380
1381 LFNQPTHCEBVDSDTOPKAPGIDDIKTLEBKLSLSESSSSGAQASVSLFTSVIES 1440
1380 LFNQPTHCEBVDSDTOPKAPGIDDIKTLEBKLSLSESSSSGAQASVSLFTSVIES 1440
1441 TVTPGIPITTAAPSKLITSTSTCLPPTNPLPGVIALPVTPVTPGVGVSTVSTTSQVK 1500
1440 TVTPGIPITTAAPSKLITSTSTCLPPTNPLPGVIALPVTPVTPGVGVSTVSTTSQVK 1500
1501 PGTAPSKPPLTKAPVLVGLPAGTLPSBQLPFPGPSITLQSQOPLDPAQRLTSLP 1560
1500 PGTAPSKPPLTKAPVLVGLPAGTLPSBQLPFPGPSITLQSQOPLDPAQRLTSLP 1560
1561 EXITVTSAGVPSMAAPTAITEAGTOPQKGVSYKEGQVILATSSGAGVFXMGRFQVSVA 1620
1560 EXITVTSAGVPSMAAPTAITEAGTOPQKGVSYKEGQVILATSSGAGVFXMGRFQVSVA 1620
1621 DGAQKEGKNKSEDAKSVHESSTESSVSSSESTLYKREPNGITIPGISVUPRESAH 1680
1620 DGAQKEGKNKSEDAKSVHESSTESSVSSSESTLYKREPNGITIPGISVUPRESAH 1680
1661 KTTSEAKSDPDGOTKGRFOVTTTANKVGRFSYKTEKIDTIDRKKEGPVAPPEMDLEQ 1740
1660 KTTSEAKSDPDGOTKGRFOVTTTANKVGRFSYKTEKIDTIDRKKEGPVAPPEMDLEQ 1740
1741 AVLPAVTPKKEKPELSEPSHLNGPSDEEAFLSRDVVDGSGSPHSQHLSKSLPSQNL 1800
1740 AVLPAVTPKKEKPELSEPSHLNGPSDEEAFLSRDVVDGSGSPHSQHLSKSLPSQNL 1800
1801 SGLSNSFNSSYMSDNESEDIEDLKLRLRDKHLKXIODLOSQKHIEESLYTLGL 1860
1800 SGLSNSFNSSYMSDNESEDIEDLKLRLRDKHLKXIODLOSQKHIEESLYTLGL 1860
1861 KVPRAVITIPPAAPISGRRRRPTKSGKSSRSSSLGKNSPOLSGNLGQSAASVHLHPQT 1920
1860 KVPRAVITIPPAAPISGRRRRPTKSGKSSRSSSLGKNSPOLSGNLGQSAASVHLHPQT 1920

QY 1921 LHPNIPESGONOLLOPLKSPSSDNLYSAFTSGAISVBSLAPGCGTSTNTVGATV 1980
 DB 1920 LHPNIPESGONOLLOPLKSPSSDNLYSAFTSGAISVBSLAPGCGTSTNTVGATV 1979
 QY 1981 NSOAAQAPPMATSRKGTFTDDHLKLVNNAARDAMNLSGRGSGHNNYGPGRARFS 2040
 DB 1980 NSOAAQAPPMATSRKGTFTDDHLKLVNNAARDAMNLSGRGSGHNNYGPGRARFS 2039
 QY 2041 APGOLCISMTSNLGGASAPISAASATSLGHFTKSMCPPOYGFPAITPFGAOWSGTGPAPO 2100
 DB 2040 APGOLCISMTSNLGGASAPISAASATSLGHFTKSMCPPOYGFPAITPFGAOWSGTGPAPO 2099
 QY 2101 PLGQFQPVGTASLQNFNINSNQKSI SNPPGSGNLT 2136
 DB 2100 PLGQFQPVGTASLQNFNINSNQKSI SNPPGSGNLT 2135
 RESULT 2
 US-10-288-798-9
 ; Sequence 9, Application US/10288798
 ; Publication No. US20030207299A1
 ; GENERAL INFORMATION:
 ; APPLICANT: BANDMAN, Olga; NGUYEN, Daniel B;
 ; APPLICANT: WALIA, Narinder K.; HARALYA, April J.A.;
 ; APPLICANT: GURURAJAN, Rajagopal; DING, Li;
 ; APPLICANT: PATTERSON, Chandra; YUE, Henry;
 ; APPLICANT: BAUGHN, Mariah R.; TRIBOLEY, Catherine M.;
 ; APPLICANT: THORNTON, Michael; ELLIOTT, Vicki S.;
 ; APPLICANT: LU, Yan; ISON, Craig H.;
 ; APPLICANT: AU-YOUNG, Janice; TANG, Y. Tom;
 ; APPLICANT: AZIMZAI, Yalda; BURRILL, John D.;
 ; APPLICANT: MARCUS, Gregory A.; ZINGLER, Kurt A.;
 ; APPLICANT: LU, Dyoung Alina M.; LAI, Preeti G.;
 ; APPLICANT: RAMKUMAR, Jayalaxmi; WARREN, Bridget A.;
 ; APPLICANT: KEARNEY, Liam; POLICKY, Jennifer L.;
 ; APPLICANT: THANHAVEU, Kavitha; BURFORD, Neil
 ; TITLE OF INVENTION: HUMAN KINASES
 ; FILE REFERENCE: PI-0209 USA
 ; CURRENT APPLICATION NUMBER: US/10/288, 798
 ; PRIOR FILING DATE: 2002-11-01
 ; PRIOR APPLICATION NUMBER: PCT/US01/27219
 ; PRIOR FILING DATE: 2001-08-31
 ; PRIOR APPLICATION NUMBER: US 60/240,542
 ; PRIOR FILING DATE: 2000-10-13
 ; PRIOR APPLICATION NUMBER: US 60/238,389
 ; PRIOR FILING DATE: 2000-10-06
 ; PRIOR APPLICATION NUMBER: US 60/236,499
 ; PRIOR FILING DATE: 2000-09-29
 ; PRIOR APPLICATION NUMBER: US 60/234,902
 ; PRIOR FILING DATE: 2000-09-22
 ; PRIOR APPLICATION NUMBER: US 60/232,654
 ; PRIOR FILING DATE: 2000-09-14
 ; PRIOR APPLICATION NUMBER: US 60/231,357
 ; PRIOR FILING DATE: 2000-09-08
 ; PRIOR APPLICATION NUMBER: US 60/229,873
 ; PRIOR FILING DATE: 2000-08-31
 ; NUMBER OF SEQ ID NOS: 48
 ; SOFTWARE: PERL Program
 ; SEQ ID NO 9
 ; LENGTH: 2135
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; OTHER INFORMATION: Incyte ID No. US20030207299A1 5502218CD1
 US-10-288-798-9
 Query Match 99.8%; Score 10794.5; DB 15; Length 2135;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 2134; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

DB 1 MSGAAEKQSTPGSLFSPAPAPKXGSSDSSVBEKLGAAADAVTGRTEYRRRRHT 60
 QY 61 MDKSRGAAATTTTTERHFFRRSVYCDSNATALLPGILPLSPDIPAAVQASAPPEH 120
 DB 61 MDKSRGAAATTTTTERHFFRRSVYCDSNATALLPGILPLSPDIPAAVQASAPPEH 120
 QY 121 REETVATATQVAAQPPAAAPGEQAVAGAPETVSTSTKDBPVQSPSTVGSKEEPP 180
 DB 121 REETVATATQVAAQPPAAAPGEQAVAGAPETVSTSTKDBPVQSPSTVGSKEEPP 180
 QY 181 ARSGGGGSAKEPEEESQOQDDIELEETKAVGNSNGRFLKFDIEIGRGSFKTYVKLD 240
 DB 181 ARSGGGGSAKEPEEESQOQDDIELEETKAVGNSNGRFLKFDIEIGRGSFKTYVKLD 240
 QY 241 TETVEAVAMCELODRKLTSEBQRFKEEBAEMKGLQHPNIVREYDSWESTVYKGCIVLV 300
 DB 241 TETVEAVAMCELODRKLTSEBQRFKEEBAEMKGLQHPNIVREYDSWESTVYKGCIVLV 300
 QY 301 TELMTSGTLTKYLRKRVYMKIKVLRSCROILKGLQFLHRTPTPIHRDLKCDNIFITGP 360
 DB 301 TELMTSGTLTKYLRKRVYMKIKVLRSCROILKGLQFLHRTPTPIHRDLKCDNIFITGP 360
 QY 361 TGSYKIGDLGLATLKRAAFKSVIGTPEFNAPEMEYEEKYDSVDVYAFGMCLMATSEY 420
 DB 361 TGSYKIGDLGLATLKRAAFKSVIGTPEFNAPEMEYEEKYDSVDVYAFGMCLMATSEY 420
 QY 421 PYSCQAAQIYRRVTSQVYAPASDKAIPBVEIIEGCIKRONKERYSIKDLINHAFO 480
 DB 421 PYSCQAAQIYRRVTSQVYAPASDKAIPBVEIIEGCIKRONKERYSIKDLINHAFO 480
 QY 481 BEGVRVEALBEDDGEKIAIKMLRLEDIKKGYKDNDAIESFLEEDVEDVAQEM 540
 DB 481 BEGVRVEALBEDDGEKIAIKMLRLEDIKKGYKDNDAIESFLEEDVEDVAQEM 540
 QY 541 VESGYCEGDHKTMAKAIKDVSLIKKREPORLVEEQQEESLQOYVEQSSASQ 600
 DB 541 VESGYCEGDHKTMAKAIKDVSLIKKREPORLVEEQQEESLQOYVEQSSASQ 600
 QY 601 TGIKOLPASTGIPYASTTSASVSTOVEPEPADHQOQOYQOPSTISVLSGTQVSDG 660
 DB 601 TGIKOLPASTGIPYASTTSASVSTOVEPEPADHQOQOYQOPSTISVLSGTQVSDG 660
 QY 661 SSVFTSRVSSQOTVSGSHQEAHSTGVPHIPSTVQAQOSPHGVYPPSSVAQGSQG 720
 DB 661 SSVFTSRVSSQOTVSGSHQEAHSTGVPHIPSTVQAQOSPHGVYPPSSVAQGSQG 720
 QY 721 QPSSSSLTGVSSQPIQHPPOQOQIGIOTAPPOQTVOYSLQSTSSSEATTAQVSOQAP 780
 DB 721 QPSSSSLTGVSSQPIQHPPOQOQIGIOTAPPOQTVOYSLQSTSSSEATTAQVSOQAP 780
 QY 781 QVLPQVSAQKOSTQVSAVAPAPVAVAOPTPTTLASSVDASHSAGSDGNNV 840
 DB 781 QVLPQVSAQKOSTQVSAVAPAPVAVAOPTPTTLASSVDASHSAGSDGNNV 840
 QY 841 PSSSGRHEGRTGRHRSKVSRSRHEKTSRPLRLTLYNSNGKDRVVECOLETHNKKMT 900
 DB 841 PSSSGRHEGRTGRHRSKVSRSRHEKTSRPLRLTLYNSNGKDRVVECOLETHNKKMT 900
 QY 900 FKFDLDGNDPELIATIMVNNDFILAIERESFVDQVREIEKXDEMSESDVSVEBEDQGL 959
 DB 900 FKFDLDGNDPELIATIMVNNDFILAIERESFVDQVREIEKXDEMSESDVSVEBEDQGL 959
 QY 961 ESLOGKDDYGFSSQQLBEFQKPIPASSMPOQIGIPSTSLQVVSAGARFVSVPEPS 1019
 DB 961 ESLOGKDDYGFSSQQLBEFQKPIPASSMPOQIGIPSTSLQVVSAGARFVSVPEPS 1019
 QY 1021 RLRESKVPFSEKTDYTAASTAOSPGMNLSHSASSLSLQAFSLRRAQMTGEGNTAPPNF 1080
 DB 1021 RLRESKVPFSEKTDYTAASTAOSPGMNLSHSASSLSLQAFSLRRAQMTGEGNTAPPNF 1080
 QY 1081 SHTGTFPVVPPFLSIAGVPTTAAATAPVATSSPNDISTVIOSEVTVPTTEGIACV 1140
 DB 1081 SHTGTFPVVPPFLSIAGVPTTAAATAPVATSSPNDISTVIOSEVTVPTTEGIACV 1140

Db 1080 SHTGPTFPVPPPLSSAGVPTTAATAAPVATSPSPNDISTVIOSEVTVPTEEGTACV 1139
 QY 1141 ATSTGVTSGLDIPVSESPVLSSVVSITITPAVVISITTSPLQVPTSTSEIIVSSIA 1200
 Db 1140 ATSTGVTSGLDIPVSESPVLSSVVSITITPAVVISITTSPLQVPTSTSEIIVSSIA 1199
 QY 1201 LYSVYVVSATSSAGSGTATPPGPPVAVVQQAAGSTTVGATLTSVSTTSFSTASOLS 1260
 Db 1200 LYSVYVVSATSSAGSGTATPPGPPVAVVQQAAGSTTVGATLTSVSTTSFSTASOLS 1259
 QY 1261 IOLSSSTSTPLAETVAVVASHSLDKTSHSSTTGLAFSLAPSSSSSGAGVSVYSIQPGG 1320
 Db 1260 IOLSSSTSTPLAETVAVVASHSLDKTSHSSTTGLAFSLAPSSSSSGAGVSVYSIQPGG 1319
 QY 1321 LHPVIVSVIASVILPQAAAGTSTPLLPVPSIPPLVQPVANVAVQOTLIHSQOPAL 1380
 Db 1320 LHPVIVSVIASVILPQAAAGTSTPLLPVPSIPPLVQPVANVAVQOTLIHSQOPAL 1379
 QY 1381 LHPVIVSVIASVILPQAAAGTSTPLLPVPSIPPLVQPVANVAVQOTLIHSQOPAL 1440
 Db 1380 LHPVIVSVIASVILPQAAAGTSTPLLPVPSIPPLVQPVANVAVQOTLIHSQOPAL 1439
 QY 1441 TVTPTGIPPTAVAPSKLITSTTSTCLPPTNLPLGTVALPVTPVVTGQVSTPVSTTSGVX 1500
 Db 1440 TVTPTGIPPTAVAPSKLITSTTSTCLPPTNLPLGTVALPVTPVVTGQVSTPVSTTSGVX 1499
 QY 1501 PGTAPSKPLTKAPVLPVGTBLPAGTLPSEGLPPFPGBSLTQSQOPLLEDDAOLRLTSLP 1560
 Db 1500 PGTAPSKPLTKAPVLPVGTBLPAGTLPSEGLPPFPGBSLTQSQOPLLEDDAOLRLTSLP 1559
 QY 1561 EXLTVTAVGVPVMAAFTATTEACTOPKGVQVKEGVPVLTATSSGAGVPMKRGVQVAA 1620
 Db 1560 EXLTVTAVGVPVMAAFTATTEACTOPKGVQVKEGVPVLTATSSGAGVPMKRGVQVAA 1619
 QY 1621 DGAQKEKKNSEDAKSVHFESESTSESVLSSSESTLVKEPEPGITIPGISDVPESAH 1680
 Db 1620 DGAQKEKKNSEDAKSVHFESESTSESVLSSSESTLVKEPEPGITIPGISDVPESAH 1679
 QY 1681 KTTASEAKSDTGQPTKGRFOVTTTANKVGRFVSKTEDKTTDTRKKGPPVAPPFMDLEQ 1740
 Db 1680 KTTASEAKSDTGQPTKGRFOVTTTANKVGRFVSKTEDKTTDTRKKGPPVAPPFMDLEQ 1739
 QY 1741 AVLPAAVLPKKEKPELSEPSHNGPSDEPAFLSRVDVDSGSPHSPHOLSKSLPSONL 1800
 Db 1740 AVLPAAVLPKKEKPELSEPSHNGPSDEPAFLSRVDVDSGSPHSPHOLSKSLPSONL 1799
 QY 1801 SOSLSNFSNSSYMSGSNDSEDIEDDLKELRLRDLKELKE1QDLOSRQKHEIESLYTKLG 1860
 Db 1800 SOSLSNFSNSSYMSGSNDSEDIEDDLKELRLRDLKELKE1QDLOSRQKHEIESLYTKLG 1859
 QY 1861 KVPAAVITPPAAPLSGRRRRPTKSGKSSRSSSLGNKSPOLSGNLSGQSAASVILHPQOT 1920
 Db 1860 KVPAAVITPPAAPLSGRRRRPTKSGKSSRSSSLGNKSPOLSGNLSGQSAASVILHPQOT 1919
 QY 1921 LHPGNTIPESONOLLOPLKSPSSSDNLYSAFTSGAISVPSLSAPGGSTSTTVGATV 1980
 Db 1920 LHPGNTIPESONOLLOPLKSPSSSDNLYSAFTSGAISVPSLSAPGGSTSTTVGATV 1979
 QY 1981 NSQAAQAPPAVMTSSRKGTFTDHLKLVNNAARDAMNLSGRRGSGKHNNYEGPGMARFS 2040
 Db 1980 NSQAAQAPPAVMTSSRKGTFTDHLKLVNNAARDAMNLSGRRGSGKHNNYEGPGMARFS 2039
 QY 2041 APGOLCISMTSNLGSAPISAAASATSLGHFTKSMCPPOQYGPAPATPPCAOMSGTGGAPQ 2100
 Db 2040 APGOLCISMTSNLGSAPISAAASATSLGHFTKSMCPPOQYGPAPATPPCAOMSGTGGAPQ 2099
 QY 2101 PLGGQFQVGTASLQNFNTSNLOKSI SNPPGSLRTT 2136
 Db 2100 PLGGQFQVGTASLQNFNTSNLOKSI SNPPGSLRTT 2135

RESULT 3
 US-10-336-472-230

; Sequence 230. Application US/10336472
 ; Publication No. US20040043929A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Anderson, David W.
 ; APPLICANT: Ballinger, Robert A.
 ; APPLICANT: Baumgartner, Jason C.
 ; APPLICANT: Burgess, Catherine E.
 ; APPLICANT: Casman, Stacie J.
 ; APPLICANT: Chant, John S.
 ; APPLICANT: Bergns, Constance
 ; APPLICANT: Gangolli, Bsha A.
 ; APPLICANT: Edinger, Shlomil R.
 ; APPLICANT: Eilerman, Karen
 ; APPLICANT: Furtak, Katarzyna
 ; APPLICANT: Gilbert, Jennifer A.
 ; APPLICANT: Gunther, Erik
 ; APPLICANT: Gorman, Linda
 ; APPLICANT: Guo, Xiaojia Sasha
 ; APPLICANT: Ji, Weizhen
 ; APPLICANT: Li, Li
 ; APPLICANT: Liu, Xiaohong
 ; APPLICANT: Miller, Charles E.
 ; APPLICANT: Millet, Isabelle
 ; APPLICANT: Padigaru, Muralidhara
 ; APPLICANT: Patrajan, Meera
 ; APPLICANT: Rastelli, Luca
 ; APPLICANT: Macdougall, John R.
 ; APPLICANT: Mishra, Vishnu
 ; APPLICANT: Pena, Carol E.A.
 ; APPLICANT: Spaderna, Steven K.
 ; APPLICANT: Shimkets, Richard A.
 ; APPLICANT: Smithson, Glenda
 ; APPLICANT: Spytek, Kimberly A.
 ; APPLICANT: Stone, David J.
 ; APPLICANT: Shenoy, Suresh G.
 ; APPLICANT: Ort, Tatiana
 ; APPLICANT: Taupier Jr, Raymond J.
 ; APPLICANT: Tchernev, Velizar T.
 ; APPLICANT: Verneet, Corine A.M.
 ; APPLICANT: Wolenc, Adam R.
 ; APPLICANT: Zerhusen, Bryan D.
 ; APPLICANT: Zhong, Mei
 ; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
 ; FILE REFERENCE: 21402-533C
 ; CURRENT FILING DATE: 2003-01-03
 ; PRIOR APPLICATION NUMBER: 09/746,491
 ; PRIOR FILING DATE: 2000-12-20
 ; PRIOR APPLICATION NUMBER: 10/005,041
 ; PRIOR FILING DATE: 2001-12-04
 ; PRIOR APPLICATION NUMBER: 10/023,681
 ; PRIOR FILING DATE: 2001-12-18
 ; PRIOR APPLICATION NUMBER: 10/024,212
 ; PRIOR FILING DATE: 2001-12-18
 ; PRIOR APPLICATION NUMBER: 10/055,569
 ; PRIOR FILING DATE: 2001-10-26
 ; PRIOR APPLICATION NUMBER: 10/080,334
 ; PRIOR FILING DATE: 2002-02-21
 ; PRIOR APPLICATION NUMBER: 10/092,900
 ; PRIOR FILING DATE: 2002-03-07
 ; PRIOR APPLICATION NUMBER: 10/136,826
 ; PRIOR FILING DATE: 2002-05-01
 ; PRIOR APPLICATION NUMBER: 10/236,417
 ; PRIOR FILING DATE: 2002-09-06
 ; PRIOR APPLICATION NUMBER: 60/345,092
 ; PRIOR FILING DATE: 2002-01-04
 ; Remaining Prior Application data removed - See File Wrapper or PALM.
 ; NUMBER OF SEQ ID NOS: 230
 ; SOFTWARE: Cytaseqdist version 0.1
 ; SEQ ID NO 230
 ; LENGTH: 2362
 ; TYPE: prt

ORGANISM: MNK1
US-10-336-472-230

Query Match 98.8%; Score 10677; DB 12; Length 2382;
Best Local Similarity 89.6%; Pred. No. 0;
Matches 2135; Conservative 0; Mismatches 1; Indels 246; Gaps 1;

QY 1 MSGGAEEQOSTPGSLFLSPAPAPKXSSSDSVGEKLGAAADAVTGTTEYYRRRHT 60
DB 1 MSGGAEEQOSTPGSLFLSPAPAPKXSSSDSVGEKLGAAADAVTGTTEYYRRRHT 60
QY 61 MDKSRGAAATTTTTEHFRFRSVCDNATALELPGPLSLPQPSITAAAPQSAPEPH 120
DB 61 MDKSRGAAATTTTTEHFRFRSVCDNATALELPGPLSLPQPSITAAAPQSAPEPH 120
QY 121 REETVATATSQVAQPPAAAAPGQAVAGPAPSTVPSTSKDRPVSOPLVSGKEEPP 180
DB 121 REETVATATSQVAQPPAAAAPGQAVAGPAPSTVPSTSKDRPVSOPLVSGKEEPP 180
QY 181 ARSGGGGSAKEPQEBERSQODDIELETKAVGMSNDRFLKPIEIGRGSFKTVYGLD 240
DB 181 ARSGGGGSAKEPQEBERSQODDIELETKAVGMSNDRFLKPIEIGRGSFKTVYGLD 240
QY 241 TETTVAYAMCELODRKLTCSRORFKERAEMLKGLQHNIVRFYDSWESTYKGCIVLV 300
DB 241 TETTVAYAMCELODRKLTCSRORFKERAEMLKGLQHNIVRFYDSWESTYKGCIVLV 300
QY 241 TETTVAYAMCELODRKLTCSRORFKERAEMLKGLQHNIVRFYDSWESTYKGCIVLV 300
DB 241 TETTVAYAMCELODRKLTCSRORFKERAEMLKGLQHNIVRFYDSWESTYKGCIVLV 300
QY 301 TELMTSGTLKTYLRFKVMKIKVLRSWCROILKGLQFLHTPTPIIHRDLKCDNI FTTPG 360
DB 301 TELMTSGTLKTYLRFKVMKIKVLRSWCROILKGLQFLHTPTPIIHRDLKCDNI FTTPG 360
QY 361 TGSYKIGDLGLATIKRASPAKSVIGTEPFMAPEMYEEKYDESVDVYAFGCMLEMATSEY 420
DB 361 TGSYKIGDLGLATIKRASPAKSVIGTEPFMAPEMYEEKYDESVDVYAFGCMLEMATSEY 420
QY 421 PYSECQAAQIYRRVTSGVKPAEDKVAIPEVKEIIIGCTRONDERYSIKDLINAFQ 480
DB 421 PYSECQAAQIYRRVTSGVKPAEDKVAIPEVKEIIIGCTRONDERYSIKDLINAFQ 480
QY 481 EETGVRELAEEDGEEKIAIKMLRIDIKKLKGYKDNALIEFSPLEBDVEDVAQEM 540
DB 481 EETGVRELAEEDGEEKIAIKMLRIDIKKLKGYKDNALIEFSPLEBDVEDVAQEM 540
QY 541 VESGYVEEGHKTMAKAIKORVSLIKKRQRQIVREEOEKKKQEBESSLKQVEQSSASQ 600
DB 541 VESGYVEEGHKTMAKAIKORVSLIKKRQRQIVREEOEKKKQEBESSLKQVEQSSASQ 600
QY 601 TGIKOLPSASTGTPLASTTSASVSTOVEPEPEADHQOIQOQPSISVLSIDGTVDSSGQ 660
DB 601 TGIKOLPSASTGTPLASTTSASVSTOVEPEPEADHQOIQOQPSISVLSIDGTVDSSGQ 660
QY 661 SSVETESRVSQQTVSYSQHEQNAHSTGTPGHIPTVQAQSQPHGVPPSSVAQGSQSG 720
DB 661 SSVETESRVSQQTVSYSQHEQNAHSTGTPGHIPTVQAQSQPHGVPPSSVAQGSQSG 720
QY 721 QBSSSSLTGVSSQPIQHPQOQOGIOQTAPPOQTVOYYSISQTSSEATTAGAPVSOQAP 780
DB 721 QBSSSSLTGVSSQPIQHPQOQOGIOQTAPPOQTVOYYSISQTSSEATTAGAPVSOQAP 780
QY 781 QVLPQVSAGKQ----- 791
DB 781 QVLPQVSAGKQ----- 791
QY 792 ----- 791
DB 792 ----- 791
QY 841 VSGIPISTPHVSTAGTGFSSLPITMAAGITQPLITLASSATTAIPGVSTVVSQPLTL 900
DB 841 VSGIPISTPHVSTAGTGFSSLPITMAAGITQPLITLASSATTAIPGVSTVVSQPLTL 900
QY 901 QPVTOLPSQVHPQLQPAVQSMGI PANTGQAAEVPLSSGDLVYQGFPPRLPQYPGDSNI 960
DB 901 QPVTOLPSQVHPQLQPAVQSMGI PANTGQAAEVPLSSGDLVYQGFPPRLPQYPGDSNI 960
QY 792 ----- 791
DB 792 ----- 791

DB 961 APPSNVAVSCIHSTVLSPMPTEVLATPGYPPVQVPEVSNLIVPMGVGQVQVQSG 1020
QY 792 -----STQGVQVAAPEPAVAVQPCATQCTTTLASSVDSASHVDSAGMS 834
DB 1021 GSIAQAPTTSSQQAVALBSTQGVQVAAPEPAVAVQPCATQCTTTLASSVDSASHVDSAGMS 1080
QY 835 DGENEVSSSGRHHGRRTTKRHYRKSYSRSRSHKETSPPKILNVNSNGDVRVECOLETH 894
DB 1081 DGENEVSSSGRHHGRRTTKRHYRKSYSRSRSHKETSPPKILNVNSNGDVRVECOLETH 1140
QY 895 NRKAVTEKFDLDGNPEIATIMVNDPILAIERESFVDQYREIIEKADMLSEDEVSEP 954
DB 1141 NRKAVTEKFDLDGNPEIATIMVNDPILAIERESFVDQYREIIEKADMLSEDEVSEP 1200
QY 955 EGDGGLSLQCKDDYGGSSGQKLEGEKOPIPASSMPOQIGIPSSILTOVYHSGRFTIV 1014
DB 1201 EGDGGLSLQCKDDYGGSSGQKLEGEKOPIPASSMPOQIGIPSSILTOVYHSGRFTIV 1260
QY 1015 SPVESRLRESKVPSEITDTVAASSTAQSPOMNLSHASSLSLQAFSELRRAQMTGPN 1074
DB 1261 SPVESRLRESKVPSEITDTVAASSTAQSPOMNLSHASSLSLQAFSELRRAQMTGPN 1320
QY 1075 TAPNFSHTGTPPEVVPFLSSIAGVPTTAATAAPVATSSPNDISTSVYQSEVTPTE 1134
DB 1321 TAPNFSHTGTPPEVVPFLSSIAGVPTTAATAAPVATSSPNDISTSVYQSEVTPTE 1380
QY 1135 EGIAGVASTGVTSGLPIPVSESPVLSVSVSIIIPAVVSTISPSIQVTSSEI 1194
DB 1381 EGIAGVASTGVTSGLPIPVSESPVLSVSVSIIIPAVVSTISPSIQVTSSEI 1440
QY 1195 VVSSTALPSEVTSATSSAGSGTATPGPPPAVVSQQAAGSTTVGATLTSVSTTSFPS 1224
DB 1441 VVSSTALPSEVTSATSSAGSGTATPGPPPAVVSQQAAGSTTVGATLTSVSTTSFPS 1500
QY 1255 TASQLSLQSSSTSTPLAETVVVSAHSADKTSHTTGLAFSLASBSSSSPAGVSSY 1314
DB 1501 TASQLSLQSSSTSTPLAETVVVSAHSADKTSHTTGLAFSLASBSSSSPAGVSSY 1560
QY 1315 ISOPGGLHPLVPSVASTPILPQAAGPTSTPLLPQVPSIPPVQPAVPAVQOITLHS 1374
DB 1561 ISOPGGLHPLVPSVASTPILPQAAGPTSTPLLPQVPSIPPVQPAVPAVQOITLHS 1620
QY 1375 QPQPALIPNPHHCEVDSDTOPKAPGIDITKLEBKARSLSSEHSSSGAQAHSVLET 1434
DB 1621 QPQPALIPNPHHCEVDSDTOPKAPGIDITKLEBKARSLSSEHSSSGAQAHSVLET 1680
QY 1435 SLVIESTVTPGIPTTAAVPSKILTSITSTCLPPTNLPLGTVALPVTVPVTPQGVSTPVST 1494
DB 1681 SLVIESTVTPGIPTTAAVPSKILTSITSTCLPPTNLPLGTVALPVTVPVTPQGVSTPVST 1740
QY 1495 TTSQVAPGTAAPSKRPPLTKAPVLPVGTLEPAGTLPSEQLPFPFGPSITQSQOPLIEDDAOL 1554
DB 1741 TTSQVAPGTAAPSKRPPLTKAPVLPVGTLEPAGTLPSEQLPFPFGPSITQSQOPLIEDDAOL 1800
QY 1555 RRTLSPEKITTAVSAPVSMAPATAITTEAGTOPQOKVSVQKEGPVATSSGAGVFXMGFR 1614
DB 1801 RRTLSPEKITTAVSAPVSMAPATAITTEAGTOPQOKVSVQKEGPVATSSGAGVFXMGFR 1860
QY 1615 QVSVAADGAQKGRKNSDEAKSVHFPSSSTSESSVSSSPESTLVKPEPNGITIPGISSD 1674
DB 1861 QVSVAADGAQKGRKNSDEAKSVHFPSSSTSESSVSSSPESTLVKPEPNGITIPGISSD 1920
QY 1675 VPESAHTTASEKSPDTGQPTVGRQVTTANKVGRFVSUKTEKITITTKKEGPAASPP 1734
DB 1921 VPESAHTTASEKSPDTGQPTVGRQVTTANKVGRFVSUKTEKITITTKKEGPAASPP 1980
QY 1735 FMDLEQAVLPAYIPKKEKPELSEPHLNGPSSDPEAFAFSRDVDSGSPHSPHOLSSGS 1794
DB 1981 FMDLEQAVLPAYIPKKEKPELSEPHLNGPSSDPEAFAFSRDVDSGSPHSPHOLSSGS 2040
QY 1795 LPSQNTSGLSNSFNSSYMSDNESEDIEDDLKLELRRLRDGHLKEIQLQSRQKHEIES 1854
DB 2041 LPSQNTSGLSNSFNSSYMSDNESEDIEDDLKLELRRLRDGHLKEIQLQSRQKHEIES 2100

QY	1855	YTTTLGAVPPAVII	PPAAPISGRRRR	PTKSKGSSSSSS	SLGKNSPOLSGNL	SGQSAASY	1914	
Db	2101	LYTKLGVPPAVII	PPAAPISGRRRR	PTKSKGSSSSSS <th>SLGKNSPOLSGNL</th> <th>SGQSAASY</th> <td>2160</td>	SLGKNSPOLSGNL	SGQSAASY	2160	
QY	1915	LHPQOOLHPGNI	DESGONQLOLP	PKPSPSSDNL	YSAFTSDGAI	SVBLSAPGGGTSSTN	1974	
Db	2161	LHPQOOLHPGNI	DESGONQLOLP	PKPSPSSDNL	YSAFTSDGAI	SVBLSAPGGGTSSTN	2220	
QY	1975	TVGATVNSQA	AOAPPAATSSR	KGFTFDLHK	LVDNMRD	AMNLSGRRGSGKHNMYGPG	2034	
Db	2221	TVGATVNSQA	AOAPPAATSSR	KGFTFDLHK	LVDNMRD	AMNLSGRRGSGKHNMYGPG	2280	
QY	2035	MARKFSAPGOL	CISMTSNL	GGSAEISA	SATSLGHT	KSMCPPOGYG	PATPFQAOMSGT	2094
Db	2281	MARKFSAPGOL	CISMTSNL	GGSAEISA	SATSLGHT	KSMCPPOGYG	PATPFQAOMSGT	2340
QY	2095	GGPAPQPLGOF	QFQPVGTASL	QNFNLSNLOKS	ISNPGSLRTT	2136		
Db	2341	GGPAPQPLGOF	QFQPVGTASL	QNFNLSNLOKS	ISNPGSLRTT	2382		

```

1      RESULT 4
2      US-10-196-935A-2
3      Sequence 2, Application US/10196935A
4      Publication No. US20030082720A1
5      GENERAL INFORMATION:
6      APPLICANT: Lifson, Richard P
7      APPLICANT: Wilson, Frederick H
8      APPLICANT: Choate, Keith
9      APPLICANT: Ishikawa, Kazuhiko
10     APPLICANT: Nelson-Williams, Carole
11     TITLE OF INVENTION: COMPOSITIONS METHODS AND KITS RELATING TO TREATING AND DIAGNOSING
12     TITLE OF INVENTION: HYPERTENSION
13     FILE REFERENCE: 044574-5113
14     CURRENT APPLICATION NUMBER: US/10/196,935A
15     CURRENT FILING DATE: 2002-10-25
16     PRIOR APPLICATION NUMBER: US 60/306,084
17     PRIOR FILING DATE: 2001-07-17
18     NUMBER OF SEQ ID NOS: 6
19     SOFTWARE: PatentIn version 3.1
20     SEQ ID NO 2
21     LENGTH: 2382
22     TYPE: PRT
23     ORGANISM: Homo sapiens
24     US-10-196-935A-2

```

Query Match	98.8%	Score 10677	DB 14	Length 2382
Best Local Similarity	89.6%	Pred. No. 0		
Matches 215; Conservative	0	Mismatches	1	Indels 246; Gaps 1

Oy	1	MSGGAEEQSSSTPGLTFLTSPAPAPAKNSSSPSSVGEKLGAAAANAAYGRTEEYRRHT	60
Db	1	MSGGAEEQSSSTPGLTFLTSPAPAPAKNSSSPSSVGEKLGAAAANAAYGRTEEYRRHT	60
Oy	61	MDKDSRGAATTTTTEHRRFFRRSVICDSNATLLELPGLLPLSPGPSIPLAAYQSAAPPEPH	120
Db	61	MDKDSRGAATTTTTEHRRFFRRSVICDSNATLLELPGLLPLSPGPSIPLAAYQSAAPPEPH	120
Oy	121	REETVATATSOVAOQPPAAAAAGHQAVAGPAPSTVPSTSKDRPVYQSPSLVGSKEEPP	180
Db	121	REETVATATSOVAOQPPAAAAAGHQAVAGPAPSTVPSTSKDRPVQSPSLVGSKEEPP	180
Oy	121	RRSGGGGSAXKRPQERSQQDDIELETKAAGMSNDGRFLKFDLEIGKSGKTYKGLD	240
Db	121	RRSGGGGSAXKRPQERSQQDDIELETKAAGMSNDGRFLKFDLEIGKSGKTYKGLD	240
Oy	161	ARRSGGGGSAXKRPQERSQQDDIELETKAAGMSNDGRFLKFDLEIGKSGKTYKGLD	240
Db	161	ARRSGGGGSAXKRPQERSQQDDIELETKAAGMSNDGRFLKFDLEIGKSGKTYKGLD	240
Oy	241	TETTYEVAWCELODKRLTKSERQREKEAEAMKGLQHPRIYAFYDYSWESTVAGKKCIYLV	300
Db	241	TETTYEVAWCELODKRLTKSERQREKEAEAMKGLQHPRIYAFYDYSWESTVAGKKCIYLV	300
Oy	301	TELMSTGLTKYLVKRPVKMKIKVLBSMCKQILKGLQPLTRTPPIIHRDKCDNFIITGP	360
Db	301	TELMSTGLTKYLVKRPVKMKIKVLBSMCKQILKGLQPLTRTPPIIHRDKCDNFIITGP	360

D	b	301	TElMTSGTlKTYlKRFRNMlKlYlRSMQJlLKGLOfMlHRlRPPIlHRlDKDNlFI	T	P	360
Q	y	361	TGSYKlGDIglATlLKRAFSKSVlGTPEFMAPMEYKlDESVDVYAFGMClEMATSY			420
D	b	361	TGSYKlGDIglATlLKRAFSKSVlGTPEFMAPMEYKlDESVDVYAFGMClEMATSEY			420
Q	y	421	PYSQCMAAOlYRRWYSGVKPASPDKVAlPEVKEllEGCIRONKDERYSlKDLlNlAFPO			480
D	b	421	PYSQCMAAOlYRRWYSGVKPASPDKVAlPEVKEllEGCIRONKDERYSlKDLlNlAFPO			480
Q	y	481	EETVRRVlAEEDDGEKlAlKlMLRlEDlKlKGXKXONAlEFSFLEHVEDVEAOEM			540
D	b	481	EETVRRVlAEEDDGEKlAlKlMLRlEDlKlKGXKXONAlEFSFLEHVEDVEAOEM			540
Q	y	541	VESGYVCEGHlTKMAKlKORVSlLKRRROROlVREOEKKYQEBSSlLQYQEGSASQ			600
D	b	541	VESGYVCEGHlTKMAKlKORVSlLKRRROROlVREOEKKYQEBSSlLQYQEGSASQ			600
Q	y	601	TGfQOLASlTSGlPAlASTSASVlQYREPEPADHQOlQYQOPSlSVlSDGTVSGGQ			660
D	b	601	TGfQOLASlTSGlPAlASTSASVlQYREPEPADHQOlQYQOPSlSVlSDGTVSGGQ			660
Q	y	661	SSVlTESRVSQQTvSYSGQHEQhASTGTVPghlPSTVQASQPhGYPPSSVAQSQSQ			720
D	b	661	SSVlTESRVSQQTvSYSGQHEQhASTGTVPghlPSTVQASQPhGYPPSSVAQSQSQ			720
Q	y	721	QPSSSlLlTGVSSSQPlQHPQOOOGlQQTAPPOQTVOYSlSQTSTSEATlTACVSPQAP			780
D	b	721	QPSSSlLlTGVSSSQPlQHPQOOOGlQQTAPPOQTVOYSlSQTSTSEATlTACVSPQAP			780
Q	y	781	QVlRQVVSAGNQ-----			791
D	b	781	QVlRQVVSAGNQ-----			791
Q	y	792	-----			791
D	b	792	-----			791
Q	y	841	VSQPlPSTPHVSTKQTGFSSlPlTMAglQPlLlTlASSATlAlEGVSTVVSQPlTLL			900
D	b	841	VSQPlPSTPHVSTKQTGFSSlPlTMAglQPlLlTlASSATlAlEGVSTVVSQPlTLL			900
Q	y	792	-----			791
D	b	792	-----			791
Q	y	901	QPlVlQlPSQVHPQlLOPAVQSMGlPAMlGOAEVlSSGDVlYQGFPlRlPQYPGDSNI			960
D	b	901	QPlVlQlPSQVHPQlLOPAVQSMGlPAMlGOAEVlSSGDVlYQGFPlRlPQYPGDSNI			960
Q	y	792	-----			791
D	b	792	-----			791
Q	y	961	APSSNVASVClHSTVlSPMPEVlATPGYPlTVQDPYVESNlLVEMGVGVQYVSQPG			1020
D	b	961	APSSNVASVClHSTVlSPMPEVlATPGYPlTVQDPYVESNlLVEMGVGVQYVSQPG			1020
Q	y	792	-----			834
D	b	792	-----			834
Q	y	1021	GSllAQAPlTSSQQAVlBSlTQVSGVAlAEFVAlVAlQAPAlQTlTlASSVDShSDVlASGMS			1080
D	b	1021	GSllAQAPlTSSQQAVlBSlTQVSGVAlAEFVAlVAlQAPAlQTlTlASSVDShSDVlASGMS			1080
Q	y	835	DGNEVNPSSSGRHEGRTTKRHlYKlSVBSRSBHEXTSPKlRlLVNSXKGRVVEQOLETH			894
D	b	835	DGNEVNPSSSGRHEGRTTKRHlYKlSVBSRSBHEXTSPKlRlLVNSXKGRVVEQOLETH			894
Q	y	1081	DGENEVPPSSGRHEGRTTKRHlYKlSVBSRSRHEKlTSPKlRlLVNSXKGRVVEQOLETH			1140
D	b	1081	DGENEVPPSSGRHEGRTTKRHlYKlSVBSRSRHEKlTSPKlRlLVNSXKGRVVEQOLETH			1140
Q	y	895	NRKXWlTlFEDlDGNPElATlMlNNPFlAlIBESVVDQVRElTEKADENlSDVVER			954
D	b	895	NRKXWlTlFEDlDGNPElATlMlNNPFlAlIBESVVDQVRElTEKADENlSDVVER			954
Q	y	1141	NRKXWlTlFEDlDGNPElATlMlNNPFlAlIBESVVDQVRElTEKADENlSDVVER			1200
D	b	1141	NRKXWlTlFEDlDGNPElATlMlNNPFlAlIBESVVDQVRElTEKADENlSDVVER			1200
Q	y	955	EGDQGLJESlQKQDlYGFSGQKLEGEFKQPlPAssMSQOlGlPTSSlTlOVVhSAGRFlV			1014
D	b	955	EGDQGLJESlQKQDlYGFSGQKLEGEFKQPlPAssMSQOlGlPTSSlTlOVVhSAGRFlV			1014
Q	y	1201	EGDQGLJESlQKQDlYGFSGQKLEGEFKQPlPAssMSQOlGlPTSSlTlOVVhSAGRFlV			1260
D	b	1201	EGDQGLJESlQKQDlYGFSGQKLEGEFKQPlPAssMSQOlGlPTSSlTlOVVhSAGRFlV			1260
Q	y	1015	SPVPEsRlRESKVPPEsITlTVAASlTASQPMNlSHSASSlSQAQFSElRRlQMTGEPN			1074
D	b	1015	SPVPEsRlRESKVPPEsITlTVAASlTASQPMNlSHSASSlSQAQFSElRRlQMTGEPN			1074
Q	y	1261	SPVPEsRlRESKVPPEsITlTVAASlTASQPMNlSHSASSlSQAQFSElRRlQMTGEPN			1320
D	b	1261	SPVPEsRlRESKVPPEsITlTVAASlTASQPMNlSHSASSlSQAQFSElRRlQMTGEPN			1320
Q	y	1075	TAPPlFHSITGTlPVPVPPFlSSlAGVPTlTAAATlPVPATSSPPNDlSTVlQSEVlVlPTE			1134
D	b	1075	TAPPlFHSITGTlPVPVPPFlSSlAGVPTlTAAATlPVPATSSPPNDlSTVlQSEVlVlPTE			1134
Q	y	1321	TAPPlFHSITGTlPVPVPPFlSSlAGVPTlTAAATlPVPATSSPPNDlSTVlQSEVlVlPTE			1380
D	b	1321	TAPPlFHSITGTlPVPVPPFlSSlAGVPTlTAAATlPVPATSSPPNDlSTVlQSEVlVlPTE			1380
Q	y	1135	BGlAAVAlSTGVVlSGGLPlPVESAPVlSSVSSlTlPAVAlSTlTSPSlQVlPTSTSKl			1194
D	b	1135				

```

QY 1195 VVSSTALVSVTVSATASAGGCTATPGKPPAVVSOQAAGSTTVGATLTVSTTTSPS 1254
Db 1441 VVSSTALVSVTVSATASAGGCTATPGKPPAVVSOQAAGSTTVGATLTVSTTTSPS 1500
QY 1255 TASQSLIOLSSSTSTPTLAETVVAHSLDKTSHSTTGLATSLAPSSSSPGAGVSSY 1314
Db 1501 TASQSLIOLSSSTSTPTLAETVVAHSLDKTSHSTTGLATSLAPSSSSPGAGVSSY 1560
QY 1315 ISQPGGLHPLVPSVIASTPILPQAAGPTSTPLPQVPSIPPLVQPVANVPVQOTLHS 1374
Db 1561 ISQPGGLHPLVPSVIASTPILPQAAGPTSTPLPQVPSIPPLVQPVANVPVQOTLHS 1620
QY 1375 QPQPLALNQPHTHCPEVSDSTQPKAPGIDIKTLEKRLSLFSEHSSGAGHVSLET 1434
Db 1621 QPQPLALNQPHTHCPEVSDSTQPKAPGIDIKTLEKRLSLFSEHSSGAGHVSLET 1680
QY 1435 SLVIESTVPGIPTTAVAAPSKLTTSTTCLPPTNLPGTVALPPTVTVTGQVSTPVST 1494
Db 1681 SLVIESTVPGIPTTAVAAPSKLTTSTTCLPPTNLPGTVALPPTVTVTGQVSTPVST 1740
QY 1495 TTSGVKGPTASKEPLTKAPVLPVGTLPAGTLPSBQLPPPGPSLTQSQPLEDLDAOL 1554
Db 1741 TTSGVKGPTASKEPLTKAPVLPVGTLPAGTLPSBQLPPPGPSLTQSQPLEDLDAOL 1800
QY 1555 RRTISPEKITYTSAVGPVMAAPTAITEAGTQPOKGVSOVKEGPVLTSSGAGVFKMGFR 1614
Db 1801 RRTISPEKITYTSAVGPVMAAPTAITEAGTQPOKGVSOVKEGPVLTSSGAGVFKMGFR 1860
QY 1615 QVSAVAOQAKQKGNKSDAKSVHPESTSSSVLSSSPSTLVKPPNITTPGLSSD 1674
Db 1861 QVSAVAOQAKQKGNKSDAKSVHPESTSSSVLSSSPSTLVKPPNITTPGLSSD 1920
QY 1675 VPBSAHKTTASSEAKSDTQOPTKVGRCVTTTANKVGRSVKTEBKITDTCKECPVASPP 1734
Db 1921 VPBSAHKTTASSEAKSDTQOPTKVGRCVTTTANKVGRSVKTEBKITDTCKECPVASPP 1980
QY 1735 FMDLEQAVLPVAVITPKKEXPELSEPSHNLNGPSSDPEAAFLSRDVDGSGSPHSQOLSKS 1794
Db 1981 FMDLEQAVLPVAVITPKKEXPELSEPSHNLNGPSSDPEAAFLSRDVDGSGSPHSQOLSKS 2040
QY 1795 LPSQNLQOSLSNSFNSSNSMSSDNDSDIEDEBLKELRLRKHKEITDLOSRQKHIES 1854
Db 2041 LPSQNLQOSLSNSFNSSNSMSSDNDSDIEDEBLKELRLRKHKEITDLOSRQKHIES 2100
QY 1855 LYTLKGKVPVAVIIPPAAPLSGRRRRPTKSKGSKSSRSSSLGKNSPOLSGNLGQOSASV 1914
Db 2101 LYTLKGKVPVAVIIPPAAPLSGRRRRPTKSKGSKSSRSSSLGKNSPOLSGNLGQOSASV 2160
QY 1915 LHPQOTLHPGNIPESSQONLLOPLKPSSSDNLYSAFSTDGAISVPSLSAPGQSTSTN 1974
Db 2161 LHPQOTLHPGNIPESSQONLLOPLKPSSSDNLYSAFSTDGAISVPSLSAPGQSTSTN 2220
QY 1975 TVGATVNSOAAQAPPMATSSRKCTFTDDHLKLVNNAARMNLGRRGSGKHNNTYGP 2034
Db 2221 TVGATVNSOAAQAPPMATSSRKCTFTDDHLKLVNNAARMNLGRRGSGKHNNTYGP 2280
QY 2035 MARKFSAPOGLCISMTSNLGGSAFISAASATSLGHTFKSMCPPOQYFPAITPFGAOMSGT 2094
Db 2281 MARKFSAPOGLCISMTSNLGGSAFISAASATSLGHTFKSMCPPOQYFPAITPFGAOMSGT 2340
QY 2095 GGPAPQPLQGFQPVGTASLQNFNISNLQKSIINPGSNLRTT 2136
Db 2341 GGPAPQPLQGFQPVGTASLQNFNISNLQKSIINPGSNLRTT 2382

```

```

RESULT 5
US-10-052-648A-40
; Sequence 40, Application US/10052648A
; Publication No. US20040005558A1
; GENERAL INFORMATION:
; APPLICANT: Anderson, David
; APPLICANT: Burgess, Catherine

```

```

; APPLICANT: Casman, Stacie
; APPLICANT: Colman, Steven
; APPLICANT: Edinger, Shlomit R.
; APPLICANT: Ellerman, Karen
; APPLICANT: Gerlach, Valerie
; APPLICANT: Gunther, Erik
; APPLICANT: Kekuda, Ramesh
; APPLICANT: MacDougall, John R.
; APPLICANT: Mehrajan, Fuad
; APPLICANT: Patnrajan, Meera
; APPLICANT: Rothenberg, Mark
; APPLICANT: Shinkets, Richard
; APPLICANT: Smithson, Glenda
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Stone, David J.
; APPLICANT: Vernet, Corine A.M.
; APPLICANT: Zethusen, Bryan D.
; TITLE OF INVENTION: PROTEINS, POLYNUCLEOTIDES ENCODING THEM AND METHODS OF
; FILE REFERENCE: 21402-250 (CURA-550)
; CURRENT FILING DATE: 2002-12-09
; PRIOR APPLICATION NUMBER: 60/262,454
; PRIOR FILING DATE: 2001-01-18
; PRIOR APPLICATION NUMBER: 60/272,920
; PRIOR FILING DATE: 2001-03-02
; PRIOR APPLICATION NUMBER: 60/284,549
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/303,229
; PRIOR FILING DATE: 2001-07-05
; PRIOR APPLICATION NUMBER: 60/262,892
; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: 60/263,605
; PRIOR FILING DATE: 2001-01-23
; PRIOR APPLICATION NUMBER: 60/269,098
; PRIOR FILING DATE: 2001-02-15
; PRIOR APPLICATION NUMBER: 60/264,159
; PRIOR FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: 60/265,517
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/271,855
; PRIOR FILING DATE: 2001-02-27
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 97
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 40
; LENGTH: 2382
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-052-648A-40

Query Match 98.8%; Score 10677; DB 15; Length 2382;
Best Local Similarity 89.6%; Pred. No. 0;
Matches 2135; Conservative 0; Mismatches 1; Indels 246; Gaps 1;

QY 1 MSGGAEEKOSTGSLPLSPAPAPKVGSSSDSVGEKLAADAADAVGTETRRRRRT 60
Db 1 MSGGAEEKOSTGSLPLSPAPAPKVGSSSDSVGEKLAADAADAVGTETRRRRRT 60
QY 61 MDKDSRGAATTTTTHRRFFRRSVICDSNATALELPPLSLPQPSIPAAVPOAPPEPH 120
Db 61 MDKDSRGAATTTTTHRRFFRRSVICDSNATALELPPLSLPQPSIPAAVPOAPPEPH 120
QY 121 REETVATATSOVAQOPPAAGQVAVGAPASTVSSSKRPVVSQSLVSGKEPP 180
Db 121 REETVATATSOVAQOPPAAGQVAVGAPASTVSSSKRPVVSQSLVSGKEPP 180
QY 181 ARSGGSGSAKEQERSQOQDDIELETKAVGNSDGRFLKPDIEIGRSGFTVYKGLD 240
Db 181 ARSGGSGSAKEQERSQOQDDIELETKAVGNSDGRFLKPDIEIGRSGFTVYKGLD 240
QY 241 TETTVAVGCELODRKLYTSERQFKEEAEMLXGLQHPNIVRFYDWSWETVKKKCTIV 300

```

Db 241 TETTVAVAMCELOQRKLTGSRQRFKEAEMLKGLQHNTVRFVDSWESTVYKGGKCVLV 300
Qy 301 TELMTSGTLKTYLXRFVVMKIKVLRSWCROILKGLQELHTRTPPIHRDLKCNITPTGP 360
Db 301 TELMTSGTLKTYLXRFVVMKIKVLRSWCROILKGLQELHTRTPPIHRDLKCNITPTGP 360
Qy 361 TGSYKIDILGATLKRAFPKSVIGTPEFMAPEMYEKEKDSVVVYAFQCMEMATSEY 420
Db 361 TGSYKIDILGATLKRAFPKSVIGTPEFMAPEMYEKEKDSVVVYAFQCMEMATSEY 420
Qy 421 PYSCQNAQAQYRRVTSVGVKSPASFDKVAIPVKEIIEGCIRQNDERYSIKDLNHAFFQ 480
Db 421 PYSCQNAQAQYRRVTSVGVKSPASFDKVAIPVKEIIEGCIRQNDERYSIKDLNHAFFQ 480
Qy 481 EETGVAVELAEEDDGEKIAIKLMLRIEDIKLKGYKDYNEAIERSFDLERVPEDVNAQEM 540
Db 481 EETGVAVELAEEDDGEKIAIKLMLRIEDIKLKGYKDYNEAIERSFDLERVPEDVNAQEM 540
Qy 541 VESGYVEGDHKTMAKAIKDRVSLIKRREOROLVREBQEKKEBESLKQOVQSSASQ 600
Db 541 VESGYVEGDHKTMAKAIKDRVSLIKRREOROLVREBQEKKEBESLKQOVQSSASQ 600
Qy 601 TGIMOLPASASTGIPASTTSASVSVOVEPEBPAQOQOLQYQOPIISVLSGTVDSGQG 660
Db 601 TGIMOLPASASTGIPASTTSASVSVOVEPEBPAQOQOLQYQOPIISVLSGTVDSGQG 660
Qy 661 SSVTETRESVSSQOVVSYSQHEQASHSTGVPGHLPSTVQASQPHGVPPSSVNAQGSQG 720
Db 661 SSVTETRESVSSQOVVSYSQHEQASHSTGVPGHLPSTVQASQPHGVPPSSVNAQGSQG 720
Qy 721 QPSSSILTVGSSSQPIQHPOQOQGIQCTAPPOQIVQYLSQTSISSKATTAQPVISQPAF 780
Db 721 QPSSSILTVGSSSQPIQHPOQOQGIQCTAPPOQIVQYLSQTSISSKATTAQPVISQPAF 780
Qy 781 QVLPQVSAQKQ----- 791
Db 781 QVLPQVSAQKQ----- 791
Qy 792 ----- 791
Db 792 ----- 791
Qy 841 VSOIPISTHVSTAQTGFSSLPITMAAGITQPLTLASSATTAAIPGVSTVPSQLPTLL 900
Db 841 VSOIPISTHVSTAQTGFSSLPITMAAGITQPLTLASSATTAAIPGVSTVPSQLPTLL 900
Qy 792 ----- 791
Db 792 ----- 791
Qy 901 QPVTQPSQVHPOLLQPAVQSMGIPANIGQAHEVLSGDVLYGCFPRRLPQYRGDSNI 960
Db 901 QPVTQPSQVHPOLLQPAVQSMGIPANIGQAHEVLSGDVLYGCFPRRLPQYRGDSNI 960
Qy 792 ----- 791
Db 792 ----- 791
Qy 961 APSNVASVCIHSTVLSPPMPTEVLATPGYFTVVQPYVESNLVPMGVGQGVQVSQPG 1020
Db 961 APSNVASVCIHSTVLSPPMPTEVLATPGYFTVVQPYVESNLVPMGVGQGVQVSQPG 1020
Qy 792 -----STGVSYVAPAEVAVAOPOATOPTTLASSYDASHDYASGMS 834
Db 792 -----STGVSYVAPAEVAVAOPOATOPTTLASSYDASHDYASGMS 834
Qy 1021 GSLAQAPTTSSQOAVLESTQGVSAVPAVAVAOPOATOPTTLASSYDASHDYASGMS 1080
Db 1021 GSLAQAPTTSSQOAVLESTQGVSAVPAVAVAOPOATOPTTLASSYDASHDYASGMS 1080
Qy 835 DGENEVPSSSGRHEGRTRKRYRKSRSRSHKESRPKRLILNVSNKGRDVECOLETH 894
Db 835 DGENEVPSSSGRHEGRTRKRYRKSRSRSHKESRPKRLILNVSNKGRDVECOLETH 894
Qy 1081 DGENEVPSSSGRHEGRTRKRYRKSRSRSHKESRPKRLILNVSNKGRDVECOLETH 1140
Db 1081 DGENEVPSSSGRHEGRTRKRYRKSRSRSHKESRPKRLILNVSNKGRDVECOLETH 1140
Qy 895 NRKAVTFKFDLDGDNPEEIIATMVNNDPILAIERBSFYDQVEIIEKADEMISEDVSEP 954
Db 895 NRKAVTFKFDLDGDNPEEIIATMVNNDPILAIERBSFYDQVEIIEKADEMISEDVSEP 954
Qy 1141 NRKAVTFKFDLDGDNPEEIIATMVNNDPILAIERBSFYDQVEIIEKADEMISEDVSEP 1200
Db 1141 NRKAVTFKFDLDGDNPEEIIATMVNNDPILAIERBSFYDQVEIIEKADEMISEDVSEP 1200
Qy 955 EGDQGLESLQKQDVGFGSGOQLBGEFQOPIPASMPQOIGIPISLLOVHSAQRRTIV 1014
Db 955 EGDQGLESLQKQDVGFGSGOQLBGEFQOPIPASMPQOIGIPISLLOVHSAQRRTIV 1014
Qy 1201 EGDQGLESLQKQDVGFGSGOQLBGEFQOPIPASMPQOIGIPISLLOVHSAQRRTIV 1260
Db 1201 EGDQGLESLQKQDVGFGSGOQLBGEFQOPIPASMPQOIGIPISLLOVHSAQRRTIV 1260
Qy 1015 SPVESRRLRESKVFESEITDVAASTAOSPGMNLSSHASISLQOAFSELRRAQMTBEPN 1074
Db 1015 SPVESRRLRESKVFESEITDVAASTAOSPGMNLSSHASISLQOAFSELRRAQMTBEPN 1074
Qy 1261 SPVESRRLRESKVFESEITDVAASTAOSPGMNLSSHASISLQOAFSELRRAQMTBEPN 1320
Db 1261 SPVESRRLRESKVFESEITDVAASTAOSPGMNLSSHASISLQOAFSELRRAQMTBEPN 1320
Qy 1075 TAPPNFSGTPTFPVVPFLSSIAQVPTTAATAVPAATSSPNDISTSVIQSEVTPTE 1134
Db 1075 TAPPNFSGTPTFPVVPFLSSIAQVPTTAATAVPAATSSPNDISTSVIQSEVTPTE 1134
Qy 1321 TAPPNFSGTPTFPVVPFLSSIAQVPTTAATAVPAATSSPNDISTSVIQSEVTPTE 1380
Db 1321 TAPPNFSGTPTFPVVPFLSSIAQVPTTAATAVPAATSSPNDISTSVIQSEVTPTE 1380

Qy 1135 EGIAGVATSGVTSGLPIPVSESPVLSVSVSITIPAVVSISTSPSLQVPTSEI 1194
Db 1381 EGIAGVATSGVTSGLPIPVSESPVLSVSVSITIPAVVSISTSPSLQVPTSEI 1440
Qy 1195 VVSSTALPVVTSVATSASAGSSTATPGKPPVAVVQQAAGSTVVGATLTVSVSTTSPS 1254
Db 1441 VVSSTALPVVTSVATSASAGSSTATPGKPPVAVVQQAAGSTVVGATLTVSVSTTSPS 1500
Qy 1255 TASQOLSTLSSSSTPPLATTVVSAHSLDKTHSSITTGALAFSLASASSSSSPAGVSSY 1314
Db 1501 TASQOLSTLSSSSTPPLATTVVSAHSLDKTHSSITTGALAFSLASASSSSSPAGVSSY 1560
Qy 1315 ISOPGGLHPLVTVSEVIASTPILPOAAGPTSTPLLPQVPSIPPLVQPVANVAVOQTLIHS 1374
Db 1561 ISOPGGLHPLVTVSEVIASTPILPOAAGPTSTPLLPQVPSIPPLVQPVANVAVOQTLIHS 1620
Qy 1375 QPOPALLPNQPHHCPEVSDTOPKAGIDIDIKTLBEKRLSFEHSSSGAQAHSVSIET 1434
Db 1621 QPOPALLPNQPHHCPEVSDTOPKAGIDIDIKTLBEKRLSFEHSSSGAQAHSVSIET 1680
Qy 1435 SLVIESVTVTGIPITTAAPBSKLTSTSTCLPPTNLPLGTVALVTVVMPGVSTPST 1494
Db 1681 SLVIESVTVTGIPITTAAPBSKLTSTSTCLPPTNLPLGTVALVTVVMPGVSTPST 1740
Qy 1495 TTSGVKPGTAPSKRPLTKAPVLPVGTLPAGTLPSBOLPPPGSILTQSQPLEDLDAOL 1554
Db 1741 TTSGVKPGTAPSKRPLTKAPVLPVGTLPAGTLPSBOLPPPGSILTQSQPLEDLDAOL 1800
Qy 1555 RRLTSPKXITVTSVAVGVSAAPATAITBAGTOPQKGVSYQVKEGVYLTSSGAGVFKMGFR 1614
Db 1801 RRLTSPKXITVTSVAVGVSAAPATAITBAGTOPQKGVSYQVKEGVYLTSSGAGVFKMGFR 1860
Qy 1615 QVSYAADGAQKEGKNKSEDAKSVFEESTSESVLSASSSPSTLVKPEPNCITTPGISSD 1674
Db 1861 QVSYAADGAQKEGKNKSEDAKSVFEESTSESVLSASSSPSTLVKPEPNCITTPGISSD 1920
Qy 1675 VPESAHKTASAEKSDPGOPTKVRFOVTTTANKVGRFVSXKTEDKITDTKKEGPVVASP 1734
Db 1921 VPESAHKTASAEKSDPGOPTKVRFOVTTTANKVGRFVSXKTEDKITDTKKEGPVVASP 1980
Qy 1735 FMDLEQAVLPAVDPKKEPELSEPSHNGPSSDPEAFLSHDVDDGSGSPHSPHQLSXS 1794
Db 1981 FMDLEQAVLPAVDPKKEPELSEPSHNGPSSDPEAFLSHDVDDGSGSPHSPHQLSXS 2040
Qy 1795 LPSQNLQSLSNSNSNSNSNSNSNSNSNSNSNSNSNSNSNSNSNSNSNSNSNSNSNSNS 1854
Db 2041 LPSQNLQSLSNSNSNSNSNSNSNSNSNSNSNSNSNSNSNSNSNSNSNSNSNSNSNSNS 2100
Qy 1855 LYTILGKVPRAVITPPAAPLSGRRRRPTKSGKSSSSSILGNKSPOLSGNLSGQSAASV 1914
Db 2101 LYTILGKVPRAVITPPAAPLSGRRRRPTKSGKSSSSSILGNKSPOLSGNLSGQSAASV 2160
Qy 1915 LHPQOTLHPGNTIPESGONOLLOPLKPSPPSSDNIYSATFSGAISVPSLSAPGQTSSTN 1974
Db 2161 LHPQOTLHPGNTIPESGONOLLOPLKPSPPSSDNIYSATFSGAISVPSLSAPGQTSSTN 2220
Qy 1975 TVGATVNSQAADQAPPAATSSRKGFITDDLKLVDNWRDMMNLSGRGSGGHNMYEPG 2034
Db 2221 TVGATVNSQAADQAPPAATSSRKGFITDDLKLVDNWRDMMNLSGRGSGGHNMYEPG 2280
Qy 2035 MARKFSAPGOLCISMTSNLGGAPISASATSLGHTXSMCPQOYGPATPFGAOWSGT 2094
Db 2281 MARKFSAPGOLCISMTSNLGGAPISASATSLGHTXSMCPQOYGPATPFGAOWSGT 2340
Qy 2095 GGPAPOPLGOFQPVGTASLQNFNINSLQKSIINPPGSLRTT 2136
Db 2341 GGPAPOPLGOFQPVGTASLQNFNINSLQKSIINPPGSLRTT 2382

RESULT 6
US-10-408-765A-1404
; Sequence 1404, Application US/10408765A

Publication No. US20040101874A1
 GENERAL INFORMATION:
 APPLICANT: Ghosh, Soumitra S.
 APPLICANT: Fahy, Boia D.
 APPLICANT: Zhang, Bing
 APPLICANT: Gibson, Bradford W.
 APPLICANT: Taylor, Steven W.
 APPLICANT: Glenn, Gary M.
 APPLICANT: Marnock, Dale E.
 TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
 TITLE OF INVENTION: IDENTIFIED IN THE MITOCHONDRIAL PROTEOME
 FILE REFERENCE: 660088.465
 CURRENT APPLICATION NUMBER: US/10/408,765A
 NUMBER OF SEQ ID NOS: 3077
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 1404
 LENGTH: 2382
 TYPE: PR
 ORGANISM: Homo sapiens
 US-10-408-765A-1404

Query Match 98.8%; Score 10677; DB 16; Length 2382;
 Best Local Similarity 89.6%; Pred. No. 0;
 Matches 2135; Conservative 0; Mismatches 1; Indels 246; Gaps 1;

QY 1 MSGAAEKOSSTPSSLSFPAPAPKNGSSSDSSVGEKLGAAADAVTGRTEYRRRRHT 60
 Db 1 MSGAAEKOSSTPSSLSFPAPAPKNGSSSDSSVGEKLGAAADAVTGRTEYRRRRHT 60
 QY 61 MDKSRGAAATTTTTHRRFRFRSVICDSNATLPLGIPSLPPSPSPAPVPGSAPPEPH 120
 Db 61 MDKSRGAAATTTTTHRRFRFRSVICDSNATLPLGIPSLPPSPSPAPVPGSAPPEPH 120
 QY 121 REETVATATTSQVAAQPPAAAPGEQAVAPAPSTVSTSKNDPVGOPSLVSGKEPP 180
 Db 121 REETVATATTSQVAAQPPAAAPGEQAVAPAPSTVSTSKNDPVGOPSLVSGKEPP 180
 QY 181 ARSSGGGSAKEPEEERSQOQDIEELETKAVGMSNDGRFLKPDIEIGRGSFKTVYKGLD 240
 Db 181 ARSSGGGSAKEPEEERSQOQDIEELETKAVGMSNDGRFLKPDIEIGRGSFKTVYKGLD 240
 QY 241 TETVEVAMCELOORUKLTKEBROCFKEAEMLKGLQHPNIVRPFDSNESTVYKGCIVLV 300
 Db 241 TETVEVAMCELOORUKLTKEBROCFKEAEMLKGLQHPNIVRPFDSNESTVYKGCIVLV 300
 QY 301 TELMTSGTLKTYLKRFFVMKIKVLRSCROIKGLQPLHTRTPPIIHRDLKCONIFITGP 360
 Db 301 TELMTSGTLKTYLKRFFVMKIKVLRSCROIKGLQPLHTRTPPIIHRDLKCONIFITGP 360
 QY 361 TGSVKIGDGLATLTKRASFAKSVIGTPEFMAPMEYEKXDESVDVYAFGCMLEMATSEY 420
 Db 361 TGSVKIGDGLATLTKRASFAKSVIGTPEFMAPMEYEKXDESVDVYAFGCMLEMATSEY 420
 QY 421 PYSECQAAQIYRRVYSGVSPASFDKVALPEVKEIIEGCTRONKDEKSYIKDLINAFQ 480
 Db 421 PYSECQAAQIYRRVYSGVSPASFDKVALPEVKEIIEGCTRONKDEKSYIKDLINAFQ 480
 QY 481 EETGVVELAEEDGKEIAIKMLRIEDIKLKGAKXKNDNAIFPSFDEBVEDVAQEN 540
 Db 481 EETGVVELAEEDGKEIAIKMLRIEDIKLKGAKXKNDNAIFPSFDEBVEDVAQEN 540
 QY 541 VESGVYCEGDHKTMAKAIKDRVSLIRKKEQROQVLEBQEKKESSLSKQVJEGSSAQ 600
 Db 541 VESGVYCEGDHKTMAKAIKDRVSLIRKKEQROQVLEBQEKKESSLSKQVJEGSSAQ 600
 QY 601 TGIKOLPSASTGTPTASTSASVSTOVEBEPADHQOLQYQPSISVLSDGTVDSSGQ 660
 Db 601 TGIKOLPSASTGTPTASTSASVSTOVEBEPADHQOLQYQPSISVLSDGTVDSSGQ 660
 QY 661 SSVFTESRVSSQQTYSQSOHQEQAHSCTGVPGHISTVQAOSOPHGVPPSSVAQOSQ 720
 Db 661 SSVFTESRVSSQQTYSQSOHQEQAHSCTGVPGHISTVQAOSOPHGVPPSSVAQOSQ 720

QY 721 QSSSLTGVSSQPIQHPOQOQGIQOTAPPOQTQVSLQSTSSSEATTAPQVQOPAP 780
 Db 721 QSSSLTGVSSQPIQHPOQOQGIQOTAPPOQTQVSLQSTSSSEATTAPQVQOPAP 780
 QY 781 QVLPQVSAGKO----- 791
 Db 781 QVLPQVSAGKQLPVSPVPTIQGEPIQVATOPSVVPVHSGAHLFVGCPLPTPLLPQY 840
 QY 792 ----- 791
 Db 841 VSGPISTPHVSTAQTGSSSLPTMAAGITQPLTLASSATTAAIPGVSTVPSQLPTLL 900
 QY 792 ----- 791
 Db 901 QPVTQLPSQVHQLQPAVQSMGIPANLQAAEVLSSGDLVYGFPPRLPPQYGSNI 960
 QY 792 ----- 791
 Db 961 APSSNVASVCIHSTVLSPPMPTVLTATGPTVQVQVYESNLLVPMGCVGQVQVSOQ 1020
 QY 792 -----STQGVQVADAEVVAQAQATQPTTLASSVDSASHSDVAGMS 834
 Db 1021 GSIAQAEPTSSQAVLSTQGVQVADAEVVAQAQATQPTTLASSVDSASHSDVAGMS 1080
 QY 835 DGENVPSSSGRHGGRTTKRHRYKSVRSRHEKTSRPKLILNVSKNGRVVCELETH 894
 Db 1081 DGENVPSSSGRHGGRTTKRHRYKSVRSRHEKTSRPKLILNVSKNGRVVCELETH 1140
 QY 895 NRKMTKFDLDGDNPEIATIMVNDPILAIERESFVDQVREIIEKADMLSDVSEY 954
 Db 1141 NRKMTKFDLDGDNPEIATIMVNDPILAIERESFVDQVREIIEKADMLSDVSEY 1200
 QY 955 EGDQGLSLGKXDDYGGSSQKLEGEFKQPIPASSMPOQIGIPTSLTOVHNSAGRPITV 1014
 Db 1201 EGDQGLSLGKXDDYGGSSQKLEGEFKQPIPASSMPOQIGIPTSLTOVHNSAGRPITV 1260
 QY 1015 SPVPESLRSEKYPSEITPTVAASTAOFGMULSHASSLSLQQAISELRRAOMTGCN 1074
 Db 1261 SPVPESLRSEKYPSEITPTVAASTAOFGMULSHASSLSLQQAISELRRAOMTGCN 1320
 QY 1075 TAPNFSHTGPTPPVPPFLSSIAGVPTTAATAVPATSSPNDISTSVIOSEVTPTE 1134
 Db 1321 TAPNFSHTGPTPPVPPFLSSIAGVPTTAATAVPATSSPNDISTSVIOSEVTPTE 1180
 QY 1135 EGIAGVATSTGVTSGGLPIPVSESPVLSVSSITIPAVASISTTSPSLQVPTSTSEI 1194
 Db 1381 EGIAGVATSTGVTSGGLPIPVSESPVLSVSSITIPAVASISTTSPSLQVPTSTSEI 1440
 QY 1195 VVSSALYPSVTVSATSASAGSTATPPKPPAVVQOAAAGSTTVGATLTSVSTTSPS 1254
 Db 1441 VVSSALYPSVTVSATSASAGSTATPPKPPAVVQOAAAGSTTVGATLTSVSTTSPS 1500
 QY 1255 TASQLSLOLSSSTPTLAEVTVVASHSLDKTSHSSTGLAFSLAPSSSSSGAGVSSY 1314
 Db 1501 TASQLSLOLSSSTPTLAEVTVVASHSLDKTSHSSTGLAFSLAPSSSSSGAGVSSY 1560
 QY 1315 ISOPGHLVLVPSVASTPILPOAAGPTSTPLLPVPSIPPLVQVAVANVAVAOQTLIHS 1374
 Db 1561 ISOPGHLVLVPSVASTPILPOAAGPTSTPLLPVPSIPPLVQVAVANVAVAOQTLIHS 1620
 QY 1375 QPQALLPQPHTHCEVNDSDOPKAPGIDDKTEEKLRSLFSEHSSSGAQAHSVSLST 1434
 Db 1621 QPQALLPQPHTHCEVNDSDOPKAPGIDDKTEEKLRSLFSEHSSSGAQAHSVSLST 1680
 QY 1435 SLVIESTVTPGIPPTAVAPSKLLSTSTCLPPTMLPLGTVALPVVPTVPGQVSTPVST 1494
 Db 1681 SLVIESTVTPGIPPTAVAPSKLLSTSTCLPPTMLPLGTVALPVVPTVPGQVSTPVST 1740
 QY 1495 TTSGVKPGTAPSKPPLTKAPVLPVGTETLPAGTLPSBOLPPFPGPSITQSQPLLEDLDAQ 1554
 Db 1741 TTSGVKPGTAPSKPPLTKAPVLPVGTETLPAGTLPSBOLPPFPGPSITQSQPLLEDLDAQ 1800

QY 658 GQSSVTFESRVSQOQTVSYSGOHEOAHSTGTVPCHIPS2VOAOSQPHGVYPPSSVAGQ 717
 DB 659 GQSSVTFESRVSQOQTVSYSGOHEOAHSTGTVPCHIPS2VOAOSQPHGVYPPSSVAGQ 718
 QY 718 SQGQSSSLTGVSSSQPIQHPOQOQGIQOAPPOQVYVLSQSTSTSEKTTAQVSP 777
 DB 719 MOGQP-SSSLAGVSSQPVQHP-QOQGIQPTVPQOAVOYSLPQAASSSEQ-TVOQPSQ- 774
 QY 778 QAPVLPQVSAQKO-STQGVSOVAPAEVAVAOQATQPTLASSVDSASHDVAGSMDG 836
 DB 775 -----EQVSAQSTOSSTQGVSAAPPEBQTP-ITQSQPQPVPLVSSVDSASHDVAGSMDG 828
 QY 837 NENVPSSSGRHEGTTKRRHKYSVRSRSHKTSRPKLRILNVSNKGDVVECOLETHNR 896
 DB 829 NENVPSSSGRHEGTTKRRHKYSVRSRSHKTSRPKLRILNVSNKGDVVECOLETHNR 888
 QY 897 KMTVFKFDLQDNDNEELATIMVNDLILATERESFVNOVRIIEKABMLSEDEVBEQ 956
 DB 889 KMTVFKFDLQDNDNEELATIMVNDLILATERESFVNOVRIIEKABMLSEDEVBEQ 948
 QY 957 DOGLSLQKQDYGFSGSKLEGEFKOPIPASMPQOIGIPTSSLTOVHSAGRRFTVSP 1016
 DB 949 DOGLSLQKQDYGFSGSKLEGEFKOPIPASMPQOIGIPTSSLTOVHSAGRRFTVSP 1008
 QY 1017 VPBRLESEKVPBEITDTVAASTAGSPGMNLSSHASLSLQAFSELRAQMTGENT 1076
 DB 1009 VPBRLESEKVPBEITDTVAASTAGSPGMNLSSHASLSLQAFSELRAQMTGENT 1068
 QY 1077 PPNESHGPTFPVVPPELSSLAGVPTTAAT--APVATSSPPNDISTSYQSVTYPT 1134
 DB 1069 PPNESHGPTFPVVPPELSSLAGVPTTAAT--APVATSSPPNDISTSYQSVTYPT 1125
 QY 1135 EGIAGVATSTGVVSGGLPIPVSESPVLSVSVSITPIPAVSISSITSSPSLQVPTSEI 1194
 DB 1126 KGIQGVTTSGVVAAGGLTLLSVSEETLSAVSSSTAPAVVYVTTSGOVQAFSSG-- 1183
 QY 1195 VVSTALYBVTVAISASAGSSTAPGPKPRAVVSQQAAGSTTVGATLTVSTTSSFPS 1254
 DB 1184 IASTGSPGFTSTTGTGVSSVAVAPNAKPTVLLQOVAGNTAGVAVITVSTTTFEPA 1243
 QY 1255 TAOQLSLOASSSTPLATTVVVAASHLDTKSHSSSTGLAFSLAPSSSSSGAGVSY 1314
 DB 1244 MASQPSLPSSSTAPLATTVVVAASHLDTKSHSSSTGLAFSLAPSSSSSGAGVSY 1303
 QY 1315 ISQPGGLHPLVPSVIASTETILPQAAGPTSTPLLPVPSIPPLVOPANVPAVOQLTHS 1374
 DB 1304 VSGPGIHPVIVISSALASTTVLPQAVFTSTPLLPVPSIPPLVOPANVPAVOQLTHS 1363
 QY 1375 QPOPALLPNOPTHCHCEVSDTOPKAPGIDIDIKLEBKLSLPSHSSSGAQAHSVLE 1434
 DB 1364 QPOPALLPNOPTHCHCEVSDTOPKAPGIDIDIKLEBKLSLPSHSSSGAQAHSVLE 1423
 QY 1435 SYVIESVTPGIPPTTAVAPSKLLTSTSTSTCLPPTNPLPGIVALPVYVTPGVSTP-- 1491
 DB 1424 PLVAV-TVPFGIPTTAAABSKMTSTSTSTCLPPTNPLPGIVAGPVMVGVSTPSTGTH 1482
 QY 1492 ---VSTTSGVKGTAAPSKPPLTKAVLPVGTLPAGTLPSPQLPFPFGPSLQSOQPLE 1548
 DB 1483 ASAPASTATGAKPCTTPPKPSLTKTVPPVGTGLSAGTVPCBCLPFPFGPSLQSOQPLE 1542
 QY 1549 DLDQALRRTLSPEXITVTSVAVGVMAAPTAITEAGTOPQKGVSYKEGVLATSSGAV 1608
 DB 1543 DLDQALRRTLSPEXITVTSVAVGVMAAPTAITEAGTOPQKGVSYKEGVLATSSGAV 1599
 QY 1609 FKGGRQVAVADGAKGKNSKEDAKSVHFESESTESSVSSSESTLVKPEPNCIT 1668
 DB 1600 VKMGRQVAVADGAKGKNSKEDAKSVHFESESTESSVSSSESTLVKPEPNCIT 1659
 QY 1669 PGJSSDVPESAHKTASAKSDTGPQTKVGFQVTTANKVGFSSXTEDKTDKPKKG 1728
 DB 1660 SGJSLDVPSTHRTTPPEAKSGTGTQXGRFQVTTANKVGFSSXTEDKTDKPKKG 1719

QY 1729 PVASPPMDLEQAVLPVAPIPAKEKPELSEPSHLNGPSSDPEAPFLSRVDGSGSPHSP 1788
 DB 1720 PVASPPMDLEQAVLPVAPIPAKEKPELSEPSHLNGPSSDPEAPFLSRVDGSGSPHSP 1778
 QY 1789 QLSKSLPSONLSGSLNSFNSSYSSSDNESDIEDDKLELRRLRDKHLEIOLQSRQ 1848
 DB 1779 HLCSKSLPSONLSGSLNSFNSSYSSSDNESDIEDDKLELRRLRDKHLEIOLQSRQ 1838
 QY 1849 KHEIESLYTLGKVPVAVIIPPAFLSGRRRRPRTKSGSKSSRSSSLGKNSPQJSGNISG 1908
 DB 1839 KHEIESLYTLGKVPVAVIIPPAFLSGRRRRPRTKSGSKSSRSSSLGKNSPQJSGNISG 1898
 QY 1909 QSAASVHPOOTLHPNIPESGONOLQPKSPSSDNYSAFTSGAISVPELSAPQ 1968
 DB 1899 QSGSVANPQOTLHPNIPESGONOLQPKSPSSDNYSAFTSGAISVPELSAPQ 1958
 QY 1969 GTSSTNTVGAIVNSQAQAOPPAWTSRKTFTDHLKLVDMARDAMNLSGRGSGXHM 2028
 DB 1959 GTSSTNTVGAIVNSQAQAOPPAWTSRKTFTDHLKLVDMARDAMNLSGRGSGXHM 2018
 QY 2029 NYEGPGARKFSAPOUCISNTSLGSSAPISAASATSLGFTYSMCPPOQYGPAPFG 2088
 DB 2019 NYEGPGARKFSAPOUCISNTSLGSSAPISAASATSLGFTYSMCPPOQYGPAPFG 2078
 QY 2089 AOWSGTGPAPOPLQGFQVGTASLQNFNISNLOKSTSNPPGSNLRKT 2136
 DB 2079 TOWSGTGPAPOPLQGFQVGTASLQNFNISNLOKSTSNPPGSNLRKT 2126

RESULT 8
 US-10-052-648A-38
 ; Sequence 38, Application US/10052648A
 ; Publication No. US2004000558A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Anderson, David
 ; APPLICANT: Burgess, Catherine
 ; APPLICANT: Casman, Stacie
 ; APPLICANT: Colman, Steven
 ; APPLICANT: Edinger, Shlomit R.
 ; APPLICANT: Elferman, Karen
 ; APPLICANT: Gerlach, Valerie
 ; APPLICANT: Gunther, Erik
 ; APPLICANT: Kerkuda, Ramesh
 ; APPLICANT: MacDougall, John R.
 ; APPLICANT: Mehraban, Fuad
 ; APPLICANT: Raturajan, Meera
 ; APPLICANT: Rothenberg, Mark
 ; APPLICANT: Shinkets, Richard
 ; APPLICANT: Smithson, Glenda
 ; APPLICANT: Spylek, Kimberly A.
 ; APPLICANT: Stone, David J.
 ; APPLICANT: Verneet, Corine A.M.
 ; APPLICANT: Zernusen, Bryan D.
 ; TITLE OF INVENTION: PROTEINS, POLYNUCLEOTIDES ENCODING THEM AND METHODS OF
 ; FILE REFERENCE: 21402-250 (CURA-550)
 ; CURRENT APPLICATION NUMBER: US/10/052,648A
 ; PRIOR FILING DATE: 2002-12-09
 ; PRIOR APPLICATION NUMBER: 60/262,454
 ; PRIOR FILING DATE: 2001-01-18
 ; PRIOR APPLICATION NUMBER: 60/272,920
 ; PRIOR FILING DATE: 2001-03-02
 ; PRIOR APPLICATION NUMBER: 60/284,549
 ; PRIOR FILING DATE: 2001-04-18
 ; PRIOR APPLICATION NUMBER: 60/303,229
 ; PRIOR FILING DATE: 2001-07-05
 ; PRIOR APPLICATION NUMBER: 60/262,892
 ; PRIOR FILING DATE: 2001-01-19
 ; PRIOR APPLICATION NUMBER: 60/263,605
 ; PRIOR FILING DATE: 2001-01-23
 ; PRIOR APPLICATION NUMBER: 60/269,098
 ; PRIOR FILING DATE: 2001-02-15
 ; PRIOR APPLICATION NUMBER: 60/264,159

; PRIOR FILING DATE: 2001-01-25
 ; PRIOR APPLICATION NUMBER: 60/265,517
 ; PRIOR FILING DATE: 2001-01-31
 ; PRIOR APPLICATION NUMBER: 60/271,855
 ; PRIOR FILING DATE: 2001-02-27
 ; Remaining Prior Application data removed - See file Wrapper or PALM.
 ; NUMBER OF SEQ ID NOS: 97
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 38
 ; LENGTH: 670
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-052-648A-38

Query Match 31.5%; Score 3408; DB 15; Length 670;
 Best Local Similarity 100.0%; Pred. No. 1.3e-162;
 Matches 668; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MSGGAEEKOSTGSLPLSPAPAPKNGSSSDSVGEKLGAAADAVTGRTEYRRRRHT	60
DB	1	MSGGAEEKOSTGSLPLSPAPAPKNGSSSDSVGEKLGAAADAVTGRTEYRRRRHT	60
QY	61	MDXDSRGAATTTTTHRRFRRSVTCDSNATALELPLPLPDPSPAPAPPEPH	120
DB	61	MDXDSRGAATTTTTHRRFRRSVTCDSNATALELPLPLPDPSPAPAPPEPH	120
QY	121	REETVATATNSQVAAQPPAAAPGEQAVAGPAPSTVSSSKDRPVSGPSLVSKEPP	180
DB	121	REETVATATNSQVAAQPPAAAPGEQAVAGPAPSTVSSSKDRPVSGPSLVSKEPP	180
QY	181	ARSGSGGSAKPEPEERSSQDDIEELJETAAGVNSNGRFLKPDIEGRSFKTVYGLD	240
DB	181	ARSGSGGSAKPEPEERSSQDDIEELJETAAGVNSNGRFLKPDIEGRSFKTVYGLD	240
QY	241	TEETVEAVAMCELQDRKLTSEKRFKEAEMLKGLQHPNIVRFYDSWESTVKGKCIIVL	300
DB	241	TEETVEAVAMCELQDRKLTSEKRFKEAEMLKGLQHPNIVRFYDSWESTVKGKCIIVL	300
QY	301	TEMTSGTLKTYLKRFRVMKIVLRSMCRQILKGLQHLHTPTPIIHRDLKCDNIFITGP	360
DB	301	TEMTSGTLKTYLKRFRVMKIVLRSMCRQILKGLQHLHTPTPIIHRDLKCDNIFITGP	360
QY	361	TGSVKIGDLGATLKRASTFAKSVIGTEPFMAPPEMYEKEKYDESUVVAFGCMLEMATSEY	420
DB	361	TGSVKIGDLGATLKRASTFAKSVIGTEPFMAPPEMYEKEKYDESUVVAFGCMLEMATSEY	420
QY	421	PYSECQAAQIYRRVTSQVAPASPDXYAIPVKEIIEGCIRQNDERSYIKDLINHAFFQ	480
DB	421	PYSECQAAQIYRRVTSQVAPASPDXYAIPVKEIIEGCIRQNDERSYIKDLINHAFFQ	480
QY	481	BEIGVAVELAEEDGEEKIATKLMRIEDIKLKGGKYDNEAIEFSFLDERVPEDVAQEM	540
DB	481	BEIGVAVELAEEDGEEKIATKLMRIEDIKLKGGKYDNEAIEFSFLDERVPEDVAQEM	540
QY	541	VESGVYECGDHKTAKAIKDRVSLIKRRRQRLVREOEKKOEESLIKOOVESASQ	600
DB	541	VESGVYECGDHKTAKAIKDRVSLIKRRRQRLVREOEKKOEESLIKOOVESASQ	600
QY	601	TGIRQLPSASTGIPASTTSASVSTQVPEPEADQHOOLQYQOPSSIVLSDGTVDGQ	660
DB	601	TGIRQLPSASTGIPASTTSASVSTQVPEPEADQHOOLQYQOPSSIVLSDGTVDGQ	660
QY	661	SSVFTESR 668	
DB	661	SSVFTESR 668	

RESULT 9
 US-10-460-545-2
 ; Sequence 2, Application US/10460545
 ; Publication No. US20040005624A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Millennium Pharmaceuticals, Inc.

; APPLICANT: Tayber, Olga
 ; TITLE OF INVENTION: 84573, A Human Protein Kinase Family
 ; FILE REFERENCE: MP102-080PLM
 ; CURRENT APPLICATION NUMBER: US/10/460,545
 ; CURRENT FILING DATE: 2003-06-12
 ; PRIOR APPLICATION NUMBER: US/60/388,031
 ; PRIOR FILING DATE: 2002-06-12
 ; NUMBER OF SEQ ID NOS: 7
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 2
 ; LENGTH: 1743
 ; TYPE: PRT
 ; ORGANISM: Homo Sapiens
 US-10-460-545-2

Query Match 22.6%; Score 2439; DB 15; Length 1743;
 Best Local Similarity 33.6%; Pred. No. 1.7e-113;
 Matches 746; Conservative 246; Mismatches 617; Indels 610; Gaps 80;

QY	27	NGSSSDSVGEKLGAAADAVTGRTEYRRRRHTMOKDSRGAATTTTTHRRFRRSV	86
DB	18	DGTSFERNVPQ---VAATLTVEARKX-----KNSTFSASGETVERKFRKSV--	63
QY	87	DSNATALELPLSLPDPSPAPAPPEPHREETVATATNSQVAAQPPAAAPGEQ	146
DB	64	-----EMTEDQKVAESSP-----KORIKAAANNIPRVKLPNSVLKGGQ	102
QY	147	AVAGPAPSTVSSSKDRPVSGPSLVSKEPPAPASGSGSAKPEPEERSSQDDIE	206
DB	103	EYVEEQSKSTSEISXD-----CFKENEKEME-----EE	132
QY	207	LETKAGVNSNGRFLKPDIEGRSFKTVYKGLDTEETVEAVAMCELQDRKLTSEKRF	266
DB	133	AEKAAVATSSSGRFLKPDIEGRSFKTVYKGLDTEETVEAVAMCELQDRKLTSEKRF	132
QY	267	EEAEMLKGLQHPNIVRFYDSWESTVKGKCIIVLVTIMTSGLTKYLRKFRVMKIVLRS	326
DB	193	EEAEMLKGLQHPNIVRFYDSWESTVKGKCIIVLVTIMTSGLTKYLRKFRVMKIVLRS	252
QY	327	WCRQILKGLQHLHTPTPIIHRDLKCDNIFITGPVSGVKIGDLGATLKRASTFAKSVIGT	386
DB	253	WCRQILKGLQHLHTPTPIIHRDLKCDNIFITGPVSGVKIGDLGATLKRASTFAKSVIGT	312
QY	387	PEFAPPEMYEKEKYDESUVVAFGCMLEMATSEYSECQAAQIYRRVTSQVAPASPD	446
DB	313	PEFAPPEMYEKEKYDESUVVAFGCMLEMATSEYSECQAAQIYRRVTSQVAPASPD	372
QY	447	VALPEVEKIIEGCIRQNDERSYIKDLINHAFFQEBETGVAVELAEEDD--GEKTAIKLWL	504
DB	373	VTPDEVEKIIEGCIRQNDERSYIKDLINHAFFQEBETGVAVELAEEDD--GEKTAIKLWL	431
QY	505	RIDIKLKGGKYDNEAIEFSFLDERVPEDVAQEMVSGGVYECGDHKTAKAIKDRVSL	564
DB	432	-VEDPKLKGGKYDNEAIEFSFLDERVPEDVAQEMVSGGVYECGDHKTAKAIKDRVSL	490
QY	565	IKRRRQRLVREOEKKOEESLIKOOVESASQGTGIRQLPSASTGIPASTTSASVS	624
DB	491	IKRRRQRLVREOEKKOEESLIKOOVESASQGTGIRQLPSASTGIPASTTSASVS	533
QY	625	TQVPEPEADQHOOLQYQOPSSIVLSDGTVDGQSSVFTSRVSSQOTVSYGS	679
DB	534	TGACCEFEVDQHOOLQYQOPSSIVLSDGTVDGQSSVFTSRVSSQOTVSYGS	591
QY	680	QHEQASTGIVPGHISTV-QAQ--SQHGVYPPSVAAQSQSQSQSSSLTGVSSQPIQ	737
DB	592	NQ-----TMSQVANSINPQAEVNVPGQIT-----SSQQLVIG	622
QY	738	HPQGGIGIQTAPPOQVYVSLSTSTSSSEATTAQVYVSLPQVLPVVSAGKOSTQVGS	797
DB	623	HYQVVSGLQKHSKLTQPIILPLVQO-----STVLPV-HVIGPVPVQPOVSLPLTVQKVP	676
QY	798	QVAP-AEPVAVAPQATQPTTIASSVDSVSHSDVAGSGSDGNNV-----PSSSGRHBGRT	851

Db 677 QIKPVSQVGAEQOALLKPDL---VSLNDQVAT---TKENSSPDNPGNKORIK 729
 QY 852 TKRHYKRSVRSRSHKTSRPRKILNLSNKGDRVBCOLETHNRKMTFFEDLDGNE 911
 Db 730 QRR-----ASCPEKOTKQVLVLYQVSTSGDNKVEQLTTHNKMTFFEDVDGDAPE 783
 QY 912 ELATIMVNDLTLAIERESFVDQVREIEKADMLSEDSVPEPGDQLESLQKDDYGF 971
 Db 784 DIAYWEDNFWLESEKKEFEELRAIVGAQOELI--HVHATERATGVDSITVDSNSQ 841
 QY 972 SGQKLESEFKQPIIPASSMPOQIGIPSSULQVHSHAGRRITVSPVESRLRESKVPSE 1031
 Db 842 TGSS-----BOVOINSTQTSNESAPQS---SPVGNMFC----- 874
 QY 1032 ITDVAASIAQSPGMNLSHSASSLQAFSELRAQCTEGEPAPMFHSTGTPFPVP 1091
 Db 875 INQIIRRETIOS-----PFSLOH----- 892
 QY 1092 PFLSSINGVPTTAATAVPATSPPN-DISTSVIGSEVTVPTREGIAGVATSTGVVTSQ 1150
 Db 893 ---SMSAVP---GRHPLPSPKNTSNKEISRDLTLIENNPCHRALFTSKSEKDDYVDG 944
 QY 1151 GLPIPVSESPVLSVSSITIPAVGISTSPSLQVPTISELV---VSTALYBSVTY 1207
 Db 945 ---KISECASVETKQPIALYQVEDNRQIMAPVNTSSYSTSV 984
 QY 1208 SATSASAGSTATPPKPPAVNSQQAAGSTVGATLTSVSTTSPSTASOLSTQSSST 1267
 Db 985 RAYVAEEGLTKQASIFIPVPCHQ-----TASQADALMHPRESSTQ----- 1026
 QY 1268 STPLTAETVVVAHSLDKTSHSSTTGLAFSLAPSSSSSPAGVSVYISQPGGLHPLVP 1327
 Db 1027 ---TSGNSLTTLAFDQKQPTLSVQCPAMDAAERISQEG-----E 1061
 QY 1328 SVLASTPLQPAAGPTSTPLLPQVSPPLVQPPAVNPVQUTLIHQOPALHPQPH 1387
 Db 1062 TVNTAASPKVTIPIPTGPLE-----PTLQ-----T 1090
 QY 1388 HCEVDSDTOPKAGIDD-IKTLLEKLS--LFSHSSSGAQHASVLSLSTVLESTVTPG 1445
 Db 1091 TVLESDEBERPKLEFADNRKITDEKRNLIYQHSIS-----SIYESQ----- 1135
 QY 1446 IPTTAAPSKLITSTSTCLPPTNLPGLVALPVTPVT---PQGVSTPVST---TTS 1497
 Db 1136 ---KDTOSIDSPFSSABDTLSCPTEVIALSHGICKOSPOSNFQGTGS 1183
 QY 1498 GVKPCTAPSKRP-----LYKAPVLPVGTLPBAGLPSB-QLP--PFGPSTLQGS 1543
 Db 1184 KTLISNVAASOPANISVFKRDLNVITVP-----SELCLHMSSDASLPDGPFAVPAVSS 1238
 QY 1544 QOPLDLDAQLRRLTSPKXITVTSVAVPVSMAPTAITEAGTQPOKQSVKQGVPLATS 1603
 Db 1239 ---GAIHQVTEHEMSATAP---DPILTRSTADTRALNCK----- 1279
 QY 1604 SGAGVFKMGFOYVAADGAQKKGKSEDAKSVHESSSTSSVLS-SSSPSTLVKPE 1662
 Db 1280 AMGSGFORGFQVITTPQ--QOSAKMTSFGIEHISVSEFTHSSSEAFITANSQVLE 1337
 QY 1663 PNGITLPGISSDVPEAHKTTASEAKSDTQGTQKVGFOVTTANKVGRSVKTEDKIT 1722
 Db 1338 P-----ATQNPKTSPSYEKLQ-AIQ 1356
 QY 1723 DTKKEG---PVASPPMDLEQAV---LPVAVPKKKEPE-----LSEPSH 1761
 Db 1357 ETCKENKGVKQGDNLFSACETDVSSVTPKEKEFHHSATSSMOGSELKKREIL 1416
 QY 1762 NG---PSSDPE-AAFLSRDVEDSG-----SPHS---PHOLSKSLPSONLSQSLNS 1807
 Db 1417 TACKOPSSSEFASILA---GSGKSVAKTGPESSQCLPHH-----EBOVAAQTOSS 1465
 QY 1808 F--NSSVMSDNDSDIEDEDLKELRLRDKHKEIODOSRQKHEIESIYTGKGVPPA 1865

Db 1466 FYSPSSPMSSDSEIEDEDLKVELQRLREKHIOEVNLTQCNKEIQELTERLSIKDS 1525
 QY 1466 -----VTPPAFLSGRRRPRPTSKSGSKSSRSSLSLNKSPQSLNSGSAABVHPQ 1919
 Db 1526 KTQSTELPLPABP-----RPNSEFKSLNRS-----POS 1555
 QY 1920 TLHPGNIIPESGQNLQPLKPSFSSDNLXSAFTSDGALSVPSLAPQGTSTNTVAT 1979
 Db 1556 LTH-----VDN-----GLVADPLCIVE 1572
 QY 1980 VNSQAQAQPPAMTSSKGTFTDHLKLVNWARDAN-----NLGRRGSKGHMVEG 2032
 Db 1573 SNAASCOQSP-----ASKKGMFTDHLKLVDMTKEAVGNSLIRKPSLNQKOSQKLFEN 1628
 QY 2033 PGMAKFSAPQOL-----CISMNSNLGSAPISSAATSLGHFTKSMCPPOQYGP 2083
 Db 1629 MNKYSB-NTPSTNGYISTWISSLSQIRGAVPTSLPQSLSPSFGP---PLSSYMPHVCO 1684
 QY 2084 ---ATPFGAOWSGTGGPAPQPL-----GOFQPVGTASLQWFINISLQKSTSNPPG 2130
 Db 1685 YNVAAGGYVQWVGISGTTQGSVVIPOAGSGCFQP---GNMQAFPTSSVQNPATIFPG 1741
 RESULT 10
 US-10-433-794-17
 ; Sequence 17, Application US/10433794
 ; Publication No. US20040077044A1
 ; GENERAL INFORMATION:
 ; APPLICANT: YUE, Henry; DING, Li;
 ; APPLICANT: LAB, Ereeti G.; GRIFFIN, Jennifer A.;
 ; APPLICANT: GURURAJAN, Rajagopal; BAUGHN, Mariah R.;
 ; APPLICANT: ISON, Craig H.; RAMKUNAR, Jayalakmi;
 ; APPLICANT: TRIBOULEY, Catherine M.; SWARNAKAR, Anita;
 ; APPLICANT: BURFORD, Neil; BANDMAN, Olga;
 ; APPLICANT: THORNTON, Michael; KHAN, Farrah A.;
 ; APPLICANT: WALIA, Narinder K.; NGUYEN, Daniel B.;
 ; APPLICANT: ELIOTT, Vicki S.; XU, Yuming;
 ; APPLICANT: LU, Yan; HAPALIA, April J.A.;
 ; APPLICANT: YAO, Monique G.; GANDHI, Ameena R.;
 ; APPLICANT: ARVIZU, Chandra S.; FORSTHE, Ian J.
 ; TITLE OF INVENTION: KINASES AND PHOSPHATASES
 ; FILE REFERENCE: PI-0311 USN
 ; CURRENT APPLICATION NUMBER: US/10/433,794
 ; PRIOR FILING DATE: 2003-06-04
 ; PRIOR APPLICATION NUMBER: PCT/US01/47431
 ; PRIOR FILING DATE: 2001-12-04
 ; PRIOR APPLICATION NUMBER: US 60/254,034
 ; PRIOR FILING DATE: 2000-12-06
 ; PRIOR APPLICATION NUMBER: US 60/251,814
 ; PRIOR FILING DATE: 2000-12-07
 ; PRIOR APPLICATION NUMBER: US 60/255,756
 ; PRIOR FILING DATE: 2000-12-14
 ; PRIOR APPLICATION NUMBER: US 60/256,172
 ; PRIOR FILING DATE: 2000-12-15
 ; PRIOR APPLICATION NUMBER: US 60/257,416
 ; PRIOR FILING DATE: 2000-12-22
 ; PRIOR APPLICATION NUMBER: US 60/260,912
 ; PRIOR FILING DATE: 2001-01-10
 ; PRIOR APPLICATION NUMBER: US 60/264,644
 ; PRIOR FILING DATE: 2001-01-25
 ; PRIOR APPLICATION NUMBER: US 60/266,017
 ; PRIOR FILING DATE: 2001-02-02
 ; NUMBER OF SEQ ID NOS: 40
 ; SOFTWARE: PERL Program
 ; SEQ ID NO 17
 ; LENGTH: 1345
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; OTHER INFORMATION: Incyte ID No: 7482044CD1
 US-10-433-794-17

Query Match 19.7%; Score 2130.5; DB 16; Length 1345;
 Best Local Similarity 38.8%; Pred. No. 3.6e-98; Indels 475; Gaps 36;
 Matches 537; Conservative 114; Mismatches 258; Indexes 475; Gaps 36;
 Db RFRFRVCDNSNATALEPLSLPSPSIPAAVPGSAPPEPHREET-----VTAT 128
 78 RFRFRVCDNSNATALEPLSLPSPSIPAAVPGSAPPEPHREET-----VTAT 128
 26 RFRFRVCDNSNATALEPLSLPSPSIPAAVPGSAPPEPHREET-----VTAT 128
 129 ATGQAQAQPPAA-----AAGEQAQVAPASSTVSSSKRPVSQPSLVSKKEPP-- 180
 75 ARGRPAPAPAPAAVAPGAPGAPADAGPEP-----VGTQEPGDPPTAAAVETAPAPDGG 128
 181 -----ARSGGGGSAK-EP-----QEESSOQODDIEELETAKVMSNDGRFLKDE 226
 129 PREBAATVRKEDEGAARPEPGRTRRDEPEEDEDUL--KAVATSIDGRFLKDE 186
 227 IGRSFKTVYKGLDTEETVAVMCELQDRKLTSEKRFKEEAMKGLQHPNIVREYDS 286
 187 IGRSFKTVYKGLDTEETVAVMCELQDRKLTSEKRFKEEAMKGLQHPNIVREYDS 246
 287 WESTVWKKCIIVVTEMTSGTLKTYLKRKWKIKYLRBWCQIILKGLQFHTRPPI 346
 247 WESSAKKRCIIVVTEMTSGTLKTYLKRKWKIKYLRBWCQIILKGLQFHTRPPI 306
 347 HRDLKCNIFITGPTGSVKIGDLGLATLKRASFASVIGTPEFAPAPMEEEKYDESVDVY 406
 307 HRDLKCNIFITGPTGSVKIGDLGLATLKRASFASVIGTPEFAPAPMEEEKYDESVDVY 366
 407 AFQCMLEMTSEYSEYSECONAAQIYRVTSQVSPASFDKVAIPEVKEIIEGCIQNKDE 466
 367 AFQCMLEMTSEYSEYSECONAAQIYRVTSQVSPASFDKVAIPEVKEIIEGCIQNKDE 426
 467 RYSLKDLINAFPEELGVEVLAEEDEGK--IAITLMRIEIKLKQKYNKMEIEF 524
 427 RYSLKDLINAFPEELGVEVLAEEDEGK--IAITLMRIEIKLKQKYNKMEIEF 484
 525 SFDERDVPEDVAOEMESGVVCGDHKTAKAIKDREVSILIKRREORQVREEOEKKKQ 584
 485 TFLDEKTPBEVAOEMESGVVCGDHKTAKAIKDREVSILIKRREORQVREEOEKKKQ 542
 585 EESSLKQOVDSASQGTGIRKQIPASSTGIPASTSASVSTOV-----EPPEEADQHQ 638
 543 DVGS-----PKKANGPPVPLQVTVYHAQAQPGPPEPEEADQHQ 583
 639 QLOVQOQSI-----SVLSDGVDSGQSSSVFTESRSVQGVVSGSQHEQAHSYG----- 688
 584 ---LLPPLTPSATSLASDSTFDSGQGVSTVSDQ--SSQGSVMLGSLADAPSPAQCVCS 639
 689 ---TVPGHLPSTVOAQSGPHG-----VYPP 710
 640 PVSEGEVLPQSLPS--LGAYQQPTAPAPGLPVGSVPAPACPSLQOHPPDPAWSPAPVLP 698
 711 SSVAQGSQGSQPS-----SSSLTGV 730
 699 PSTMPGPGQAPAPGQAPPPPLAQPTPLPVLAQPVPLQPVPLPPLPPLPPLPPLPPLP 758
 731 SSSEPIQHPPQO--QGIQO-----TAPQOQTVQ 756
 759 AQKPLQMPQAPLQPLAQVPPQMPPIPVVPITPLAGIDGLPPLPDLPLTATVPMPPPQ 818
 757 Y----- 757
 819 YFSFPAVILPSLAALPPLPASAPLPLQAVLPHRPGAPLAMPCKTIVNAPATIPLLAVAP 878
 758 -----SLQSTSSSATTAAQVVSQPO----- 778
 879 GVALSIHSAVAQUPGQPVVPAFAFPQMAPTDVPSPHNTVQNMBAATPQAPLPPQPTLPP 938
 779 ----- 778
 939 QPVLPPQPTLPPQPVLPQPTLPPQPVLPQPVLPQPVLPQPVLPQPVLPQPVLPQPVLP 998
 779 -----APQVLPQ----- 785

Db 999 AAPAATPGSIIILGHAPFAVAVDAQVPTVPVPPAAVLSPLPEVLLPAPAPELLPPQFPSS 1058
 786 ---VSAGKOS--TQVVSQVAP-----EPVAVAQPOATQPTTLA 819
 1059 LATVASASVQSPQTATLLPPLPAPNPPLPGPGIASPQPTVOLTVPEPQOEBSAQDKPPGPE 1118
 820 SSVDS--AHSVAVAG--MSQGENEVSPSSSGRHEGRTTKRHRKSVRSRREKTSRPLRI 876
 1119 QCESTGSDVTSCKELSDSCGA--FGGRLEGRARAKHRSTRARSQERARSRLTI 1177
 877 LANSNKGDRVVECOLETHNRKMTFFKFDLDGNDPBEIATIMVANDFLALERESFVDQVR 936
 1178 LANCNGDKMVEQQLFETHNKHVTFKFDLDGADAPDEIATVMEHDFILOARETFIRQMK 1237
 937 EITKXDEMISEVSVPEBDDGLE-----SLQKDDVFGSGQKLEGEFQOPIPASSMPQ 992
 1238 DVMDKXEDMLSEPTDADSGSDPSTSPHLSCTGIG--TGESSROSQANAPV-----YQ 1288
 993 QIGIPSSILQVHASGRRETVSPVPSRLESKVPFSEITDTVAASTAOSPGMNLSSHA 1052
 1289 Q-----NVHTGKRWITICPVAEHAPFA-----PSSPPLPLSSLPCLPFLRMSCA 1335
 1053 SLS 1056
 1336 SVLA 1339

RESULT 11
 US-10-052-648A-37
 ; Sequence 37, Application US/10052648A
 ; Publication No. US2004000558A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Anderson, David
 ; APPLICANT: Burgess, Catherine
 ; APPLICANT: Casman, Stacie
 ; APPLICANT: Coleman, Steven
 ; APPLICANT: Edinger, Shlomit R.
 ; APPLICANT: Ellerman, Karen
 ; APPLICANT: Gerlach, Valerie
 ; APPLICANT: Gunther, Erik
 ; APPLICANT: Kekuda, Ramesh
 ; APPLICANT: Macdougall, John R.
 ; APPLICANT: Mehraban, Fuad
 ; APPLICANT: Patturajan, Meera
 ; APPLICANT: Shimkets, Richard
 ; APPLICANT: Smitheon, Glenda
 ; APPLICANT: Spytek, Kimberly A.
 ; APPLICANT: Stone, David J.
 ; APPLICANT: Vernet, Corine A.M.
 ; APPLICANT: Zephusen, Bryan D.
 ; TITLE OF INVENTION: USING THE SAME
 ; FILE REFERENCE: 21402-250 (CURA-550)
 ; CURRENT APPLICATION NUMBER: US/10/052,648A
 ; CURRENT FILING DATE: 2002-12-09
 ; PRIOR APPLICATION NUMBER: 60/262,454
 ; PRIOR FILING DATE: 2001-01-18
 ; PRIOR APPLICATION NUMBER: 60/272,920
 ; PRIOR FILING DATE: 2001-03-02
 ; PRIOR APPLICATION NUMBER: 60/284,549
 ; PRIOR FILING DATE: 2001-04-18
 ; PRIOR APPLICATION NUMBER: 60/303,229
 ; PRIOR FILING DATE: 2001-07-05
 ; PRIOR APPLICATION NUMBER: 60/262,892
 ; PRIOR FILING DATE: 2001-01-19
 ; PRIOR APPLICATION NUMBER: 60/263,605
 ; PRIOR FILING DATE: 2001-01-23
 ; PRIOR APPLICATION NUMBER: 60/269,098
 ; PRIOR FILING DATE: 2001-02-15
 ; PRIOR APPLICATION NUMBER: 60/264,159
 ; PRIOR FILING DATE: 2001-01-25


```

; PRIOR APPLICATION NUMBER: 60/265,517
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/271,855
; PRIOR FILING DATE: 2001-02-27
; Remaining Prior Application data removed - See File Wrapper or PAM.
; NUMBER OF SEQ ID NOS: 97
; SOFTWARE: Patent Ver. 2.1
; SEQ ID NO: 37
; LENGTH: 1231
; TYPE: PR
; ORGANISM: Homo sapiens
; US-10-052-648A-37

Query Match      19.7%; Score 2126; DB 15; Length 1231;
Best Local Similarity 31.9%; Pred. No. 5,4e-98;
Matches 612; Conservative 191; Mismatches 393; Indels 724; Gaps 58;

QY 10 SSTPGLFLSPAPAPKXSSSSSVGEKLGAAADAVTGTEYRRRRRMTMDKSRGAA 69
DB 2 SQTADLALRPPPLGTAGQ-----PRLGPP-----RRARRSGK----- 37
QY 70 ATTTTTHRRPFRSVIDCSNATLELPLSLPQPSIPAAVPOSAPPEPHEETVATA 129
DB 38 AEPFRSSRLSRKSSVLDGLSSWSLPASP--ADPPDP--PDSAGGPAP----- 84
QY 130 TSQVAAQPPAABAEQAVAGAPSTVPSSTKDRPVASQPSLVGSKKEPPPARSGSGGS 189
DB 85 -----SPPPSKEPEGTWTEGAP--VXAEDSARPELPDSAVPGSRP----- 127
QY 190 AKEPQOE---ERSQODDIELELTAVGMSNDGRFLKPDIEIGRSFKTVKGLDTETTE 246
DB 128 LRVEAVALERRRRQEEKEDMETQAVATSPGRYIKFEDIEIGRSFKTVVYGLDITDTE 187
QY 247 VAMELODRKLTKEBGRFKEAEMLKGLCHPNIVRFDSMESVYKGCIVLTETMTS 306
DB 188 VAMELODRKLTKEBGRFKEAEMLKGLCHPNIVRFDSMESVYKGCIVLTETMTS 247
QY 307 GTLTYLRFKVMKIKVLRWCROILKGLQFLATRTPEIIRHDKCDNIFITPGTSYKI 366
DB 248 GTLTYLRFKVMKIKVLRWCROILKGLQFLATRTPEIIRHDKCDNIFITPGTSYKI 307
QY 367 GDLGLATLKASFAKSVIGTPEFMAPEMYEKEKYDESVDVYAFGCMTEMATSEYPSYEQ 426
DB 308 GDLGLATLKASFAKSVIGTPEFMAPEMYEKEKYDEADVYAFGCMTEMATSEYPSYEQ 367
QY 427 NAAQIVRVYSGVAPSPDKVAIPEVKEIEGCIROMDERYSIKDLNHFQEEYGV 486
DB 368 NAAQIVRVYSGVAPSPDKVAIPEVKEIEGCIROMDERYSIKDLNHFQEEYGV 427
QY 487 VELAEDDGEKIALIKLWIRIEDIKKLGKYKUNBAIESFDERVDEVEDVAQEVESGYV 546
DB 428 VELAEDDGEKIALIKLWIRIEDIKKLGKYKUNBAIESFDERVDEVEDVAQEVESGYV 546
QY 547 CEGDHKTAKAIKDRVSLIKRKEQORQVREEQEKKKQEESSLKQVEQSSASQGTIKOL 606
DB 487 CEAQYQVAVARVRVAIQRKREKLRKARE-----LEAL 521
QY 607 PSAETGPTASTASVSTQY---EPEPEADQHQLOXYQPSI--SVLSDGTVSGQSS 662
DB 522 P-PEPPPPATVMAFGPPSVFPPEPEPEADQHQPLFRHASYSTTSDCTGYLSS 580
QY 663 VFTSRVSSQQTVAQVSGQHEQAHSTGTVPCHI PSTVQAQSQPHGVPPSSVAQCSQSQGP 722
DB 581 GFLDASDPAQAP-----PGGVPSL--ASHCIPSA----- 610
QY 723 SSSSLTVGVSSQPIQHPOQGGIQQTAPQQTVOYSLSTSTSEATTAAQVPSQPAQV 782
DB 611 -----FALS----- 614
QY 783 LPQVSAGKOSTQGVQVAPAPVAVAPQATOPTTLASVDSASHDVSGMDGNENYPS 842
DB 615 IPRSGPSSDFSPG-----DSVYASDAASGLSDVGEQM--G 646

```

```

QY 843 SSGRHEGRTTKRYKRSVRSRSHKTSRPKILNLVSNKGRVVECOLETHNRKMYTEK 902
DB 647 QMRPPGRNRR-----RPRSR-----LWTVSDQNDHVEECQLQTHNSKMYTFR 692
QY 903 FDLQGNPERIATIMVNDIILAIERESFVDQREIIEKADMEISEDVVSPBEDQGLFS 962
DB 693 FDLQGNPERIATIMVNDIILAIERESFVDQREIIEKADMEISEDVVSPBEDQGLFS 962
QY 963 LOGKDVGFSGQKLEGEFKQPIPASMPQOIGPTSLTQVYHSAGRFVSPESRL 1022
DB 753 QF-----EPALPALPPLPDPSEBQ-----SSTS 779
QY 1023 RESKVPSEITDVAASTAQSFGKMLSHASISLQAFSELRRAQMTGEPNAPNFSH 1082
DB 780 LEHRMTAFSTSSSPCTPLSPG--NPRSPGTPISPGIF-----PITSPP--CH 825
QY 1083 TGPTFPVPPPLSLIACVPTTAATAVAPATSSPNDISTSVIOSEVTVPEBGIAGVAT 1142
DB 826 PSPS--PFSPISSQVSNPSPHPTSPSPSSSTPE----- 859
QY 1143 STGVVTSGLPIPPVSESPVLSVYSITIPAVVISITSPSLQVPTSEIYVSTALY 1202
DB 860 -----FPV--PLSQCP-----MSLPTTSPPTSPF--CQVTLSSPFF-- 894
QY 1203 PSYVTSATASAGSTATPGPKPPAVVSGQAAGSTVGAITLTVSTTSPSTASQSLIQ 1262
DB 895 -----PPC-----PSTSPSTTAABULS 913
QY 1263 LSSSTPTPLAEVYVASHLDKTSHTTGIAFLSLAPSSSSPAGVSGSYISQPGML 1322
DB 914 LAS-----AFSLA----- 921
QY 1323 PLVPSVIASTPLIPQAAGFTSPPLPQVSPILPQVAVNPAVQOQLIHSCQPALP 1382
DB 922 -----VMTAASGL--SPSPGLS 938
QY 1383 NQPHTCPEVSDPTQAPAGIDIKYLEKILSLFSEHSSGAGHVASLETSLVISTV 1442
DB 939 QSP----- 941
QY 1443 TPGLPTTAAPSKILNTSTTCLPPTNLPLGTVALPVPVYTPQGVSTPVSTTSGYKPG 1502
DB 942 -----PA-----PSPPLP--SLPPLPPVAPGQSP--SPHTAEV-- 973
QY 1503 TAPSKPLTAPVLPVGTPLPAGTLFSEQLPPFPSPSLTQSOQLDLELDLQRLTSPX 1562
DB 974 -SEASPP-----PARLPGE----- 987
QY 1563 ITVTSAGVPVSMAPTAITAGTOPQKGVQVKEGVPVLTSSGAGVEKGRFOYVVAADG 1622
DB 988 -----ARLAPIS-----EKGKPL-----VGRFQV----- 1008
QY 1623 AQKEGKXKSEDAKSVHEESSTSESVLSSSPSTLVKPPBNGLTIIGISSDVESAHKT 1682
DB 1009 -----SSKPAEPLQPT--SPTLSSG-----PKRSTPOLTSESDTDESA-- 1048
QY 1683 TASEAKSDTQOPTVGRFOYTTANKVGRSVSKTEKIDTDTKEGVPVAPPFMDLQAV 1742
DB 1049 -----GGP-----ETRALAESR----- 1063
QY 1743 LPAPVTPKKEKREISEPSHLNGSSDPEAFLSRDVDGSGSPH--SPHOLSSKSLPSQN 1799
DB 1064 -----AAEGLAGAVEEGDDGK--EPYGVGSPQPLSHSPVWNN 1100
QY 1800 LSQSLNSPNSYVSSDNESEDIEDKLEHRLRDHKLHEIQLDQSRKHEISLTKL 1859
DB 1101 YSYS-----SLCJSSESESGEDEFWAELOSLRQHLSEVETLQTLQKKEIDLVSRLL 1155
QY 1860 GKVPAPVITPAAPALSGRRRRPTSKSGS--KSSRSSSLGNKSPQLSG-----NLSGQSAAS 1913
DB 1156 GKQPPPGIVAPAAALSSGRQR--LSKSGPITSRRNSIQREBPFGIMRRNRSLSGS7GS 1213

```

RESULT 12
 US-10-196-935A-4
 : Sequence 4, Application US/10196935A
 : Publication No. US20030082720A1
 : GENERAL INFORMATION:
 : APPLICANT: Lifton, Richard P
 : APPLICANT: Wilson, Frederick H
 : APPLICANT: Choate, Keith
 : APPLICANT: Ishikawa, Kazuhiko
 : APPLICANT: Nelson-Williams, Carole
 : TITLE OF INVENTION: COMPOSITIONS METHODS AND KITS RELATING TO TREATING AND DIAGNOSING
 : FILE OF INVENTION: HYPERTEXTENSION
 : FILE REFERENCE: 044574-5113
 : CURRENT APPLICATION NUMBER: US/10/196,935A
 : PRIOR FILING DATE: 2002-10-25
 : PRIOR APPLICATION NUMBER: US 60/306,084
 : NUMBER OF SEQ ID NOS: 6
 : SOFTWARE: PatentIn version 3.1
 : SEQ ID NO 4
 : LENGTH: 1243
 : TYPE: PRT
 : ORGANISM: Homo sapiens
 : US-10-196-935A-4

Query Match 19.7%; Score 2126; DB 14; Length 1243;
 Best Local Similarity 31.9%; Pred. No. 5.5e-98;
 Matches 612; Conservative 191; Mismatches 393; Indels 724; Gaps 58;

10 SSTPGLFLSPAPAPAKNGSSDSSVGEKLGAAADAVTGRTEBYRRRHMTMDXSGAA 69
 14 SQTADIALRPPPLGTAGQ-----PRLGPP-----RRRRRSGK----- 49
 70 ATTTTTHRRFRREVICDSNATLELPLPLPSPSTPAVPOSAPPEPPEETVATA 129
 50 AEPKRSSRSRSRSDVLGLSSWSLPASP--ADPPDP--PDAGGPAP----- 96
 130 TSQAQQPPAAAAGEQAVAGPAPSTVSTSKDPPVSPQSLVSGKEPPPARSGSGGS 189
 97 -----SPPSSKEPPEGTWTEGAP--VKAADSARPELPDSAVGPGSREP----- 139
 190 AKERQOE--ERSQOQDDIELEETKAVGNSDREFLEKDEIGRSFKTVYKGLDETTE 246
 140 LRVEAVALERRRQEEKEDEWETAVALTPDGRYLKFDIEIGRSFKTVYKGLDITTE 199
 247 VAWCELODRKLTKEERQRFKEBAEMKGLQHPNIVRYDSDWESYVKKCIVLVTLEMTS 306
 200 VAWCELODRKLTKEERQRFKEBAEMKGLQHPNIVRYDSDWESYVKKCIVLVTLEMTS 259
 307 GTLTKYLRPFKVMKIKYLRSWCROILKLOPLHTPTPIIHRDLKCNIFITPTGSVKI 366
 260 GTLTKYLRPFKVMKIKYLRSWCROILKLOPLHTPTPIIHRDLKCNIFITPTGSVKI 319
 367 GDLGLATLKRAKSAKSVIETPEFMAPMEYBEKYDESVDVAFGKCMLEMATSEYSECQ 426
 320 GDLGLATLKRAKSAKSVIETPEFMAPMEYBEKYDESVDVAFGKCMLEMATSEYSECQ 379
 427 NAAQIVRYVSGVPAKSAKSVIETPEFMAPMEYBEKYDESVDVAFGKCMLEMATSEYSECQ 486
 380 NAAQIVRYVSGVPAKSAKSVIETPEFMAPMEYBEKYDESVDVAFGKCMLEMATSEYSECQ 439
 487 VELAEEDGEEKIAIKLWIRIDIKKLGKXYKNDNEAIEFSPDLERDVEDVAVQEVESGYV 546
 440 VELAEEDGEEKIAIKLWIRIDIKKLGKXYKNDNEAIEFSPDLERDVEDVAVQEVESGYV 498
 547 CEGDHKTMAKAIKORVSLIKRRERQOLVREDEQKKQOESLSKQOYEGSSASQGTIKOL 606
 499 CEADYQVAVARVERVAIQKREKLRAR-----LEAL 533
 607 PSASTGPTASTSASSTQV---EPREPADQHQOLQYQOPST--SVLSGTYVDSGGSS 662
 534 P-DEPGPPATVVPAPGPPSVFPPEBPEPADQHQPLFRHASYSTTSQDCEITDGYLSSS 592

663 VTESRVSSQOQTVSYSGHQEHAGHSTGVPHIIPSTVQAQSPHGVPPSSVAGQSGQSG 722
 593 GFLDASDPALQ-----PGGVPSL---ASHCLPBA----- 622
 723 SSSSLTVGSSSQPIQHPOQOQGIQOTAPRQQTQVTVYSLQSTSSSEATTAQVPSQAPQV 782
 623 -----FALS----- 626
 783 LPQVSAGKOSTQGVSAVAPAEVAVAPQATQPTTLASSVDNSAHSDVAGSGDGENVPS 842
 627 IPRGPGSDSPG-----DSYASDAAGSLDVEGM--G 658
 843 SSGRHEGRTTKRHRSKVSRSREKTSRPLRLILANSKNDRYVECOLETHNRMVTEK 902
 659 QMRPPPGNLRK-----RPRSR-----LRYSVDQNDRYVECOLTHNRMVTEK 704
 903 FDLDDNPEELIATIMVNNDFILATERESFVDQVEIIEKADENLSEDSVPEPDQGLS 962
 705 FDLDDNPEELIATIMVNNDFILATERESFVDQVEIIEKADENLSEDSVPEPDQGLS 764
 963 LQCKDYGSGSQKLBGEFKQPIYASGMPOQIGIPTSSLTQVHASGRFIVSPVESRL 1022
 765 QE-----EPAPLPALPVPLPDPNSBELO-----SSTS 791
 1023 RESKVPSEITDTVAASTAQSPGWNLSHSASSLQOAFSELRRRAQMTGPNTPAPNRS 1082
 792 LEHRSWTAFTSSSPGTPPLSPG--NPRSPTGPIPSGPTE-----DITSP--CH 837
 1083 TGPTFPVPPFLSSIAQVPTTAATAVAPATSPENDISTSVIGSEVTVPEEGIAGVAT 1142
 838 PSPS--PFSPISSQVSSNPSPHPTSLPSSSTPE----- 871
 1143 STGVYTGGLPIPVSSPVLSSVSSITIPAVVISITTSLOVPTSTEIVVSTALY 1202
 872 -----PVP--PLSOP-----WSSLPTSPPTSPPT--CSQVTLSSPFF-- 906
 1203 PSVTASATASAGGSTATPGKPPAVVSAQAAGSTVGATLTVSTTSPSTASQSLQ 1262
 907 -----PBC-----PSTSPSTTAAPLSS 925
 1263 LSSSTPTPLAETVVVSAHSLDKTSHSSTTGALFSLAPSSSSPGAGVSSYISQPGALH 1322
 926 LAS-----AFSLA----- 933
 1323 PLVPSVIASTPLIPQAGPTSTPLPQVPSIPPLVQPVANVAVQOTLIHSQOPALLP 1382
 934 -----WMTVAQSLT--SPSPGLS 950
 1383 NQPHTHCPVEDSDTPQKAPGIDIKLEKLRSLFSEHSSSGAQAASVLSLTVIESTV 1442
 951 QSP----- 953
 1443 TPGIPTTAVAPSKLSTSTSCLEPTNLPLGTVALPVPVTPGQVSTPVSTTSGVKG 1502
 954 -----PA-----SLPLPPVAPGQSP--SPHTAIVE-- 985
 1503 TABSKPILTAPVLPVGTETPAGTLPBEOPLPPPGBSLTOSQPLDLIDQALRTLSPEX 1562
 986 -SEASPP-----PARPLPGE----- 999
 1563 ITVTSAGVPVSMAPTAITBAGTOPQGVQVKBGVLATSSAGAVFKMGRFQVAVAD 1622
 1000 -----ARLAPIS-----BEGKQL-----VGRFOYT----- 1020
 1623 AQKRGKNSDASVHPESSSTSSSVLSSSPBSTVKKPEPNGITTPGISDVPEASHKT 1682
 1021 -----SKKEPAPBLPQPT--SPTLSSG-----PKRSTPQLTSSSDTDESA-- 1060
 1683 TASEAKSDTQOPTKVGKGFQVTTANKVGRSVSTBKXITDTKKEGVAAPPPMDLEQAV 1742
 1061 -----GGGP-----ETREALAESDR----- 1075
 1743 LPAVTPKKEKPELSEPHLNGPSSDPEAFLSRDVDGSGSPH---SPHQLSSKSLPSQN 1799

[illegible]

```

RESULT 13
US-09-925-301-1286
; Sequence 1286: Application US/09925301
; Patent No. US20020052308A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA106
; CURRENT APPLICATION NUMBER: US/09/925,301
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05882
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1694
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 1286
; LENGTH: 453
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (38)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (101)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (110)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (286)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-301-1286

```

Query Match	19.2%	Score 2072;	DB 9;	Length 453;
Best Local Similarity	91.8%	Pred. No. 7.6e-96;		
Matches 404; Conservative	8;	Mismatches 28;	Indels 0;	Gaps 0

```

QY      81 RRSVSDSNATLIELGCLPLSLPQPSIPAAPVPSAPPEHREELVATATSOVAQCPAA 140
Db      1 KRSVICSNAIALIELGCLPLSLPQPSIPAAPVPSAPPEHREELVATATSOVAQCPAA 60
QY      141 AAPGEGAVAGPAPSTVPSSSTSKDRPVSQPSLVSKEEPPPARSGGGGSAKEPQERSQQ 200
Db      61 AAPGEGAVAGPAPRLSPAPVAKTAQCPSLALWGAKSRRRXKVAAAAGAKKEPQERSQQ 120
QY      201 QDDIELELFYKAVGMSNDGRFLKLDIEIGRGSFTYVYKGLDTEITVEVAMCELODRKLTKS 260
Db      121 QDDIELELFYKAVGMSNDGRFLKLDIEIGRGSFTYVYKGLDTEITVEVAMCELODRKLTKS 180
QY      261 EROREFEEREMKGLQHPNIVREPYDSMESVVKKKCIVYTELMTSGTLKTYLKRFEVVK 320
Db      181 EROREFEEREMKGLQHPNIVREPYDSMESVVKKKCIVYTELMTSGTLKTYLKRFEVVK 240
QY      321 IKVLRSMWCQILKGLQFLHTRTPPIIHRDLKCNIFITGPYGSVKIGDLGLATLKRASFA 380
Db      241 IKVLRSMWCQILKGLQFLHTRTPPIIHRDLKCNIFITGPYGSVKIGDLGLATLKRASFA 300
QY      381 KSVITGTFEPAPMEYEKXDESVDVYAFGMCLMATSEVPYSECQNAACIYRRVTVSGVK 440

```

Qy	Db	Qy	Db
441	301	501	421
PAFDFKVAIPVEVEII	KSVIGTPEFMAPEMYE	KLMRLIEDIKLKGKX	KLMRLIEDIKLKGKX
EGCIRONKDERYSI	KDESDVDYAFGCMLE		
KOLNHAFFOEETG	EAATSEYVS		
GRVYLAEDDEGKAI	SECONNAQYRRVTS		
500	361	520	440
PAFDFKVAIPVEVEII	PAFDFKVAIPVEVEII		
EGCIRONKDERYSI	EGCIRONKDERYSI		
KOLNHAFFOEETG	KOLNHAFFOEETG		
GRVYLAEDDEGKAI	GRVYLAEDDEGKAI		
420	401	540	460
PAFDFKVAIPVEVEII	PAFDFKVAIPVEVEII		
EGCIRONKDERYSI	EGCIRONKDERYSI		
KOLNHAFFOEETG	KOLNHAFFOEETG		
GRVYLAEDDEGKAI	GRVYLAEDDEGKAI		
440	421	560	480
PAFDFKVAIPVEVEII	PAFDFKVAIPVEVEII		
EGCIRONKDERYSI	EGCIRONKDERYSI		
KOLNHAFFOEETG	KOLNHAFFOEETG		
GRVYLAEDDEGKAI	GRVYLAEDDEGKAI		
460	441	580	500
PAFDFKVAIPVEVEII	PAFDFKVAIPVEVEII		
EGCIRONKDERYSI	EGCIRONKDERYSI		
KOLNHAFFOEETG	KOLNHAFFOEETG		
GRVYLAEDDEGKAI	GRVYLAEDDEGKAI		
480	461	600	520
PAFDFKVAIPVEVEII	PAFDFKVAIPVEVEII		
EGCIRONKDERYSI	EGCIRONKDERYSI		
KOLNHAFFOEETG	KOLNHAFFOEETG		
GRVYLAEDDEGKAI	GRVYLAEDDEGKAI		
500	481	620	540
PAFDFKVAIPVEVEII	PAFDFKVAIPVEVEII		
EGCIRONKDERYSI	EGCIRONKDERYSI		
KOLNHAFFOEETG	KOLNHAFFOEETG		
GRVYLAEDDEGKAI	GRVYLAEDDEGKAI		
520	501	640	560
PAFDFKVAIPVEVEII	PAFDFKVAIPVEVEII		
EGCIRONKDERYSI	EGCIRONKDERYSI		
KOLNHAFFOEETG	KOLNHAFFOEETG		
GRVYLAEDDEGKAI	GRVYLAEDDEGKAI		
540	521	660	580
PAFDFKVAIPVEVEII	PAFDFKVAIPVEVEII		
EGCIRONKDERYSI	EGCIRONKDERYSI		
KOLNHAFFOEETG	KOLNHAFFOEETG		
GRVYLAEDDEGKAI	GRVYLAEDDEGKAI		
560	541	680	600
PAFDFKVAIPVEVEII	PAFDFKVAIPVEVEII		
EGCIRONKDERYSI	EGCIRONKDERYSI		
KOLNHAFFOEETG	KOLNHAFFOEETG		
GRVYLAEDDEGKAI	GRVYLAEDDEGKAI		
580	561	700	620
PAFDFKVAIPVEVEII	PAFDFKVAIPVEVEII		
EGCIRONKDERYSI	EGCIRONKDERYSI		
KOLNHAFFOEETG	KOLNHAFFOEETG		
GRVYLAEDDEGKAI	GRVYLAEDDEGKAI		
600	581	720	640
PAFDFKVAIPVEVEII	PAFDFKVAIPVEVEII		
EGCIRONKDERYSI	EGCIRONKDERYSI		
KOLNHAFFOEETG	KOLNHAFFOEETG		
GRVYLAEDDEGKAI	GRVYLAEDDEGKAI		
620	601	740	660
PAFDFKVAIPVEVEII	PAFDFKVAIPVEVEII		
EGCIRONKDERYSI	EGCIRONKDERYSI		
KOLNHAFFOEETG	KOLNHAFFOEETG		
GRVYLAEDDEGKAI	GRVYLAEDDEGKAI		
640	621	760	680
PAFDFKVAIPVEVEII	PAFDFKVAIPVEVEII		
EGCIRONKDERYSI	EGCIRONKDERYSI		
KOLNHAFFOEETG	KOLNHAFFOEETG		
GRVYLAEDDEGKAI	GRVYLAEDDEGKAI		
660	641	780	700
PAFDFKVAIPVEVEII	PAFDFKVAIPVEVEII		
EGCIRONKDERYSI	EGCIRONKDERYSI		
KOLNHAFFOEETG	KOLNHAFFOEETG		
GRVYLAEDDEGKAI	GRVYLAEDDEGKAI		
680	661	800	720
PAFDFKVAIPVEVEII	PAFDFKVAIPVEVEII		
EGCIRONKDERYSI	EGCIRONKDERYSI		
KOLNHAFFOEETG	KOLNHAFFOEETG		
GRVYLAEDDEGKAI	GRVYLAEDDEGKAI		
700	681	820	740
PAF			

```

RESULT 14
US-10-114-270--80
; Sequence 80, Application US/10114270
; Publication No. US20040030110A1
; GENERAL INFORMATION:
; APPLICANT: Guo, Xiaojia
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Miller, Charles E.
; APPLICANT: Malvankar, Uriel M.
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Patnrajan, Meera
; APPLICANT: Liu, Zhaohong
; APPLICANT: Gusev, Vladimir Y.
; APPLICANT: Li, Li
; APPLICANT: Vernet, Corine
; APPLICANT: Zernhusen, Bryan D.
; APPLICANT: Gorman, Linda
; APPLICANT: Shenoy, Suresh G.
; APPLICANT: Pena, Carol E.A.
; APPLICANT: Smithson, Glendina
; APPLICANT: Burgess, Catherine E.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Shimkets, Richard A.
; APPLICANT: Gangolli, Esna A.
; APPLICANT: Taupier Jr., Raymond J.
; APPLICANT: Casman, Stacie J.
; APPLICANT: Ji, Weizhen
; APPLICANT: Anderson, David W.
; APPLICANT: Liette, Mario W.
; APPLICANT: Rastelli, Luca
; APPLICANT: Edinger, Shlomit R.
; APPLICANT: Stone, David J.
; APPLICANT: MacDougall, John R.
; APPLICANT: Rothenberg, Mark E.
; TITLE OF INVENTION: NO. US20040030110A1el Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-322C
; CURRENT FILING DATE: US/10/114,270
; PRIOR APPLICATION NUMBER: 2002-11-27
; PRIOR FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: 60/281,136
; PRIOR FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: 60/281,863
; PRIOR FILING DATE: 2001-04-05
; PRIOR APPLICATION NUMBER: 60/281,906
; PRIOR FILING DATE: 2001-04-05
; PRIOR APPLICATION NUMBER: 60/282,020
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: 60/282,930
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: 60/282,934
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: 60/283,512
; PRIOR FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: 60/283,710
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 60/284,234
; PRIOR FILING DATE: 2001-04-17
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 470
; SEQ ID NO 80

```

LENGTH: 1251
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-10-114-270-80

Query Match 18.9%; Score 2042.5; DB 12; Length 1251;
 Best Local Similarity 30.0%; Pred. No. 8,6e-94;
 Matches 605; Conservative 180; Mismatches 430; Indels 803; Gaps 56;

10 SSTPSLSPAPAPKXGSSSDSVGKLGAAADANTGRTTEVRRRHMDKSRRA 69
 2 SQTEDALIRPPPLGTAGQ-----PRLGPP-----RRARRSGK----- 37
 70 ATTTTHERFRFRSVICDSNATALEPLPLSLPPOSIPAAVPOSAPREPREETVATA 129
 38 AEPFRSRRLSRSSVDGLGILSSMLPASP--APDPPP--PDSAGGPAR----- 84
 130 TSQVAQPPAAAPCEQAVAGAPSTVPSSTSKDPVSPSLVSGKEPPPARSGSGGS 189
 85 -----SPPSSKEPPEGTGAP--VKAADSARPELPDSAVGSGREP----- 127
 190 AKERPE--ESSOODIEELETKAVGMSNDGRFLKPDIEIGRSFKVYKGLDTETIVE 246
 128 LRVEAVALERRKEDEKEDMETQAVATSPGRYUKFIDIEIGRSFKVYKGLDTETIVE 187
 247 VAWCELO--DKLTSEKREFKEEAMKGLQHPIVAFYDSESTVKGKCIYLVEIM 304
 188 VAWCELOVTRTKLRAERQRFSEVEMKGLQHPIVAFYDSESTVKGKCIYLVEIM 247
 305 TSGTLTKYKFKYKWKIKVLSWGRQIIKGLQFLHTPTPIIHRDKCNIFFITGPGSV 364
 248 TSGTLTKYLRFRFMKPRVLOKWSROILKGLHLSRVPILHRDKCNIFFITGPGSV 307
 365 KIGLGLATLKRFAFSKSVIGTPEFMAPEMYEKEYDESVDVYARGMCMLEWATSEYPS 424
 308 KIGLGLATLKRFAFSKSVIGTPEFMAPEMYEKEYDESVDVYARGMCMLEWATSEYPS 367
 425 CONAAQIYRVYTSVGPAPSPDKVAIPEYKEIIIEGICRQNKDERYSIKOLLNHAFCOETG 484
 368 CONAAQIYRVYTSVGPAPSPDKVAIPEYKEIIIEGICRQNKDERYSIKOLLNHAFCOETG 426
 485 VRVLAIEDDEDEKAIKILMIEDIKKIKKYNKNEALIESFDLERVPEDEVAOEYVSG 544
 427 VHVLAIEDDEDEKAIKILMIEDIKKIKKYNKNEALIESFDLERVPEDEVAOEYVSG 485
 545 YVCEGDHAKMAKIKDVSILKRRKQOLVREGEKKOESESLKQOVGESSASQGIK 604
 486 LVCEADYQVAVARAVREVAALQRRKREKLRKARVAAIQRRKRLKARR-----LE 537
 605 QLPASAGTPIPTASTTASVSTOV--EPREPADQHOOLQYQPSISVLSDGTVDGSGGS 661
 538 ALP-PEBGPFPATVPMAPGPPSPVPEPEBEADQHPLEFRHASY----- 583
 662 SVTFRESVSQCIYVSYSQHEQMAHSTGTVPGHISTVQAGSQPRGVYPPSSVAGSGSQG 721
 584 -----SS 585
 722 PSSSLTGVSSSQPIQHPQOQOQIGQTAAPQQTQVQYSLSTQSTSSSEATTQAQVSPQAQ 781
 586 PSNS-----YA----- 591
 782 VLPQVSAKOSTQGVSVAPAPVAVAPQATOPTTLIASVDSAHSDVASGMSDGNENVP 841
 592 -----SNAAGLSIVGEGM- 605
 842 SSGRHEGRTTKYRKYSVSRNHEK---TSRPKLRI-LVNSNKGDRVYECOLETNNRK 897
 606 GQMRPRGRMLRRFRRLRYTSQHDEGLTFSPLLTLNKVSQONDRVVECOLQTHNSK 665
 898 MTFKFLDGDNPDEELATIMVNDPILAIRESVVDQVRETIIEKADMLSEDSVVEEGD 957
 666 MTFRFDLDGDSPEELIAANVINEFILPSEBDGFLIRIETIQVETILKRD----- 717

958 QGLSLQKDDYGFSGSQKLEGEFKQPIRASSMPQOIGIPSSSLQVYHSAGRRTVSPV 1017
 718 ----- 717
 1018 PESRLRESKVPSEITDVAASTAOSPGMNLISHASLSLQOAFSELRAQMTGPNTPAP 1077
 718 -----PLTSP 722
 1078 PNFSGTGPFPVPPFLSSINGVPTTAATAAPVATSSPPNDISVYQSEVTVPTBERGI 1137
 723 P-CHPSPS--PSPISQVSNPSPHPTSSPLPSSSTPE----- 759
 1138 AGVATSGVWTSGLPIPPVSESPVLSVSSITIPAVSISITSPSLQVPTSEIYVS 1197
 760 -----FPV-PLSQCP-----MSLPTTSPPTSP-CSIQVLS 790
 1198 STALYPSVTSANTSASAGSTATPGPPPAVVSQAAGSTVGAITLVSSTTSFPTAS 1257
 791 SPFP-----PPC-----PSTSPPTTA 808
 1258 QLSIQSSSTSTPLTARTVVSASHSLDKTSHSSTTGAFSLASASSSSPGAGVSYISQ 1317
 809 APLILSLASPSL--AVMTVAQSL--LAFS-----TSSSSPG----- 840
 1318 PGGLHPIVTSVLAETPILPQAGPTSTPLLPQVPSIPPLVQVAVYPAVQOQLIHSPQ 1377
 841 -----TLSP-----GNFSPGPIPGPIPTTSPPC-----HSPS 873
 1378 PALIPNQPHHCPEVSDTQPKAPGIDIKLEBKLSLSEHSSSGAQAHSVLETSLV 1437
 874 P-----FSPISQVSNPSPHPTSSPL 895
 1438 IESTVTEGILT-PAVASKILTSTSTCLPNTNPLGTVALPVPVTPGVSTPVTIT 1426
 896 PFSSSTPEFPVPLSQCPWSSILPTTSPPTSCQVGTAVSTYPPSGKSSAAAPKWTGA 955
 1497 SGVPEGT-----APSKPPLTKAPVLPGVTELPAGTLPSQLPPFPSPSLTOS 1543
 956 GQPPRPPLHLLBULCLETHTPLEPPE--PQVPSLSLPHYIPA--PPHSPPPLPRS 1009
 1544 QQLIEDIDAQRLRTLSPEKLTTSVAVGVSMAAPALTEAGTORQKGVSGVKGQVLIATS 1603
 1010 PQ-----IPLHPPAL----- 1020
 1604 SGAGVFYKGRFQYVVAADGAQKKGKNSEDAKSVHFPSSSTSESSVLSSSPESTLVKPER 1663
 1021 -----HFHSPPAHDSFRSHS----- 1035
 1664 NGITIPGISSDVDESAAKKTATSEAKSDTQPTKVGRQVTTTANKVGRFVSKTEDKITD 1723
 1036 --LSVPGLVS--PRLHLRL-----VGRQVITSKEPAPPLPQPTSPPLTS 1079
 1724 TKKRGVAPSPFMDLEQAVIPAVIPKKEKPELSEPHLNGPSSDPEAFLSRVDVDSGS 1783
 1080 S-----PKSTPOLT-----SESDTE-----DSAGG 1102
 1784 PHSPHQLSSKSLPSQNLSSLSNSFNSSVMSNEDIDEDLKLRLRDLKHLEKIQD 1843
 1103 P-----ERHLSVET 1112
 1844 LOSROKHELESYTKGKVPVAVIIPPAALSGRRRPTKSGS-KSSRSSSLGKNSPOL 1902
 1113 LQTLQKKEIEDLXSRGKQPPGIVAPAAMLSRQR--LSKGSFTRSRNSIQRSRPMWC 1170
 1903 SGNLSGGSAAASVLAHQOQLHPGNIPESGQNL-----LQPLKPSPSDNTYSAFTSD 1955
 1171 VPAIPATQEAFA-----GEFLEPGRRLQNHSPSALRONSSSPSHNLFITNSKP 1220
 1956 GAISVPELSAPGGTSTNTVGATVNSQAQAQAPPMNT 1993
 1221 -----PLTSATSRKMSFHT--SLGSMIFILKPMYT 1250

```

RESULT 15
US-10-415-011-17
/ Sequence 17, Application US/10415011
/ Publication No. US20040053394A1
/ GENERAL INFORMATION:
/ APPLICANT: INCYTE CORPORATION
/ APPLICANT: GURURAJAN, Rajagopal
/ APPLICANT: BAUGHN, Mariah R.
/ APPLICANT: CHAWLA, Nandinder K.
/ APPLICANT: ELLIOTT, Vicki S.
/ APPLICANT: XU, Yuning
/ APPLICANT: ARVIZU, Chandra S.
/ APPLICANT: YAO, Monique G.
/ APPLICANT: RAMKOMAR, Jayalaxmi
/ APPLICANT: DING, Li
/ APPLICANT: TANG, Y. Tom
/ APPLICANT: HAFALIA, April J.A.
/ APPLICANT: NGUYEN, Daniel B.
/ APPLICANT: GANDHI, Ameena R.
/ APPLICANT: LU, Yan
/ APPLICANT: YUE, Henry
/ APPLICANT: BURFORD, Neil
/ APPLICANT: BANDMAN, Olga
/ APPLICANT: TRIBOUTY, Catherine M.
/ APPLICANT: TAL, Preeti G.
/ APPLICANT: RECIBON, Shirley A.
/ APPLICANT: LU, Dzung Anna M.
/ APPLICANT: BOROMSKY, Mark L.
/ APPLICANT: THORNTON, Michael B.
/ APPLICANT: SWANNAKER Anita
/ APPLICANT: THANGAVELU, Kavitha
/ APPLICANT: KHAN, Farrah A.
/ APPLICANT: ISON, Craig H.
/ TITLE OF INVENTION: HUMAN KINASES
/ FILE REFERENCE: PI-0262 USN
/ CURRENT APPLICATION NUMBER: US/10/415,011
/ CURRENT FILING DATE: 2003-04-18
/ PRIOR APPLICATION NUMBER: PCT/US01/47728
/ PRIOR FILING DATE: 2001-10-20
/ PRIOR APPLICATION NUMBER: US 60/242,410
/ PRIOR FILING DATE: 2000-10-20
/ PRIOR APPLICATION NUMBER: US 60/244,068
/ PRIOR FILING DATE: 2000-10-27
/ PRIOR APPLICATION NUMBER: US 60/245,708
/ PRIOR FILING DATE: 2000-11-03
/ PRIOR APPLICATION NUMBER: US 60/247,672
/ PRIOR FILING DATE: 2000-11-09
/ PRIOR APPLICATION NUMBER: US 60/249,565
/ PRIOR FILING DATE: 2000-11-16
/ PRIOR APPLICATION NUMBER: US 60/252,730
/ PRIOR FILING DATE: 2000-11-22
/ PRIOR APPLICATION NUMBER: US 60/250,807
/ PRIOR FILING DATE: 2000-12-01
/ NUMBER OF SEQ ID NOS: 44
/ SOFTWARE: PERL Program
/ SEQ ID NO 17
/ LENGTH: 1084
/ TYPE: PRT
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: misc feature
/ OTHER INFORMATION: Incyte ID No. US20040053394A1 55009053CD1
US-10-415-011-17

```

```

Query Match 18.9%; Score 2038.5; DB 12; Length 1084;
Best Local Similarity 32.8%; Pred. No. 1,1e-93;
Matches 566; Conservative 172; Mismatches 312; Indels 673; Gaps 50;

```

```

QY 207 LETKAVMSNDGFLFKEDIEIGSGFKTVYKGLDTEETVEVANCELODRKLTSESQRQFK 266
DB 1 MEFQAVATSPDGRKTKNDIEIGSGFKTVYRGDITDTTVEVANCELOTRLSRAERQRF 60
QY 267 EBAEMKGLQHPNIVRFYDSWESTVKGKCIIVLVTETMTSGTLKTYIKRPFVKIKYLR 326

```

```

DB 61 EEVEMKGLQHPNIVRFYDSWESTVKGKCIIVLVTETMTSGTLKTYIKRPFVKIKYLR 120
QY 327 WCRQILKGLQPLHTTRPPIIHRDLKCDNIETGPTSGVIGDLGLTKLRASPAKSYIGT 386
DB 121 WSRQILKGLHFLHSRVPPIIHRDLKCDNVFITGSSGVKIGDLGLTKLRASPAKSYIGT 180
QY 387 PEFMAPEMEYEEKYDESVDYVAFGCMLEMATSEYEPSECONAAQIYRYVTSGVKPAFDK 446
DB 181 PEFMAPEMEYEEKYDEAVDYVAFGCMLEMATSEYEPSECONAAQIYRYVTSGVKPAFDK 240
QY 447 VAIPVEKEITIEGTRONKDERYSIKDLNHAFFOEETGYRVEIAEEDDGEKIAIKMLRT 506
DB 241 VKIPEVKEITIEGTRONKDERYSIKDLNHAFFOEETGYRVEIAEEDDGEKIPKILMLRM 300
QY 507 EDIKLKGYKONEATEFSFDLERDVEEDVADQMVSGYVEGDHHTMAKAIIDRVSLK 546
DB 301 EDARR-GGRPRDQALEFLFQGRDAEVAEQMVALGLVCEADYQVAVARAERYVALIQ 359
QY 567 RKRECRQLVREBEQEKKEESSLKQVEQSSASQTIKQIPASASTGIPYASTTSASVSTQ 626
DB 360 RKREKLRKARE-----LEALP-PEGPPATVPMAGPSP 393
QY 627 V--EPEPPADQHQLOQYQPSI-SVLSGTVDSQSGSVFTESVSSQQTIVSYSGQH 682
DB 394 VPEPEPEPEADQHQPLFRHASYSTSDCETGYLSSSGFUDADPALQ-----445
QY 683 QAHSTGTVPGHPTVTAQASQPHGVYPPSSVAQSQSQGPSSSLGVSSSQIPQHPOQ 742
DB 446 -----PGGVPSL--AESHLCLPSA-----463
QY 743 OGIOQTAPQOQVQVSLQSTSTSEATTAQVSPQAPQVLPQVSAKQSTQGVQVABA 802
DB 464 -----FALS-----IPSGGSPSPSP-----480
QY 803 EPVAVAPQATQPTTLASSVDSASHSVASGMDNENVPSSSGRHEGRTTKRHYKSVRS 862
DB 481 -----DSYASDAASGLSVGBGM-GQMRPPGRMIR-RR-----RP 513
QY 863 RSRHEKTSRPKLILNVSKNGDRVYECOLETHNRKVTREKFDIDGNPEIATIMYNNDF 922
DB 514 RSR-----LKVTSVSDQNDVVEVCQJQTHRSKVTREKFDIDGNPEIATIMYNNDF 565
QY 923 IIAIERESEVDQVREIIEKXADEMLSEDSVEPEGDQLESLOKDDYGFSGSQKLEGEFK 982
DB 566 IIPSEDDGFLRRREITIGVETLIKRDGTGMEABDTLSPQ-----607
QY 983 QPIPASMFQOIGIPTSSLTQVNHAGRPYVSPVESKLRBSKVPESETTIVAASIAQ 1042
DB 608 EPAFLPALVPLPDPENSEIQ-----SSTSLERHMTASTSSSGTGL 652
QY 1043 SPMGNLSHASLSLQAFSELRRAQMTGPTAPNFHSHGTGTFVVPFLSSIGVPT 1102
DB 653 SPG-NPFSPEPTIPSPPIF-----PITSPP-CHPSPS--PREFDISQVSSNRS 686
QY 1103 TAAATAPVPATSPSPNDISTSVYQSEVTVPTBERGIAGVASTGVTSGLPIPPVESPV 1162
DB 697 PHTSSPLPFSSTPE-----FPV-PLSQCP-721
QY 1163 LLSVVSITIPAVVSGISTSPSLQVPTSTSEIVSSITALVPSVTSATASAGASTATRG 1222
DB 722 -----WSSLPITSPPTSPSP-CPQVTLSSPFF-----747
QY 1223 PKPFAVVSQQAAGSTTVGATLTSTSTSPSPASQSLQSSSTSTPTLATVTVVSAHS 1282
DB 748 --PFC-----PSTSPSTTAPPLSLAS-----769
QY 1283 LDKTSHSSTTGAFSLASAPSSSSSPGAGVSVYSQPGGLHPIVPSVIASTPILPOAAGP 1342
DB 770 -----AFSLA-----774
QY 1343 TSTPLLPQVPSIPPLVQVAVNPVAVOQTLHSPQALPLPNQHTHCPEVSDTQKAPG 1402

```

```
Db 775 -----VMTVAQSL--SPSPGLSSQSP----- 794
QY 1403 IDDIKTEELKRLSFSEHSSGQAASVLETSIVESTVTPGIPTAVAPSKLLTSTTS 1462
Db 795 -----PA----- 796
QY 1463 TCUPTNLPLGTVALPVTPVTPQGVSTPVSTTSQVKGPGTAPSKPPLTKAPVLPVGTTEL 1522
Db 797 ---PSPLP---SLPLPPVAPGQGESP--SPHTAEVY--SEASPP----- 832
QY 1523 PACTLPSEQLPPFPGPSLTQSQPLEDLDAQLRRTLSPEXITVTSAVGPVSMAPTAITE 1582
Db 833 PARPLPGE-----ARLAPIS----- 847
QY 1583 AGTQPOKGVSOVKKEGVLATSSGAGVFKMGRFQVSVAAADGAQKKGKNSKSEDAKSVHFESS 1642
Db 848 -----EKGKPOL-----VGRFOVT-----SKKEPAEPLPIQPT 875
QY 1643 TSSESVLSSSPESTLVKPEPMGITIPGISDVPESA---HKTASEAKSDTGQPTKVGR 1699
Db 876 ---SPTLSGS-----PKPSTPQLTSSESDTEDSAGGGETREALAESD----- 915
QY 1700 FQYTTTANKVGRFSVSKTEDKITDTKKEGVPVAPPFMDLEQAVLPAVITPKKEKPELSEPS 1759
Db 916 ---RAABGLG-----AGVEEEG----- 929
QY 1760 HUNGSSSDPEAAFLSRDVEDDGSQSPH---SPHOLSSKLPQNLQSLSNSFNSSVMSD 1816
Db 930 -----DDGK--EPQVGSPPQPLSHSPVWMNVSYs-----SLCLSEE 965
QY 1817 NESDIEDDEDLKLELRRLRDHLKEIQDLSROKHEIESLYTKLGKVPVAVIIPPAAPLSC 1876
Db 966 SESSGEDEEFMAELQSLRQKHLSEVERTLQTLQKKEIEDLYSRIGKQPPGIVAPAAMLSS 1025
QY 1877 RRRRPTKSKGS--KSSRSSSLGNKSPQLSG-----NLGSGSAAS 1913
Db 1026 RQRR--LSKGSFPTSRRNSLQRESEPPGPGIMRRNSLSSSTGS 1066
```

Search completed: September 24, 2004, 01:25:47
Job time : 232 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd

OM protein - protein search, using sw model

Run on: September 24, 2004, 01:03:14 ; Search time 41 Seconds
(without alignments)
2689.587 Million cell updates/sec

```
Title: US-10-010-720-14
Perfect score: 10812
Sequence: 1 MSGAAEKQSGSTPGSLFSP.....NISMLQKSISNPPGSMURTT 2136
```

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

```
Minimum DB seq length: 0
Maximum DB seq length: 20000000000
```

```
Post-processing: Minimum Match 0%
                  Maximum Match 100%
                  Listing first 45 summaries
```

```
Database :
Issued Patents AA:*
1: /cgn2_6/prodata/2/iaa/5A_COMB_.dep:*
2: /cgn2_6/prodata/2/iaa/5B_COMB_.dep:*
3: /cgn2_6/prodata/2/iaa/5A_COMB_.dep:*
4: /cgn2_6/prodata/2/iaa/6B_COMB_.dep:*
5: /cgn2_6/prodata/2/iaa/PCrUS_COMB_.dep:*
6: /cgn2_6/prodata/2/iaa/backfiles1_.dep:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result	Score	Query	Length	ID	Description
No.			DB		
1	10810	100.0	2136	4	US-09-854-856-14
2	10753.5	99.5	2229	4	US-09-854-856-8
3	10677	98.8	2382	4	US-09-854-856-2
4	10646	98.5	2108	4	US-09-854-856-30
5	10589.5	97.9	2201	4	US-09-854-856-24
6	10513	97.2	2354	4	US-09-854-856-18
7	10508	97.2	2076	4	US-09-854-856-46
8	10451.5	96.7	2169	4	US-09-854-856-40
9	10375	96.0	2322	4	US-09-854-856-34
10	10344	95.7	2048	4	US-09-854-856-62
11	10287.5	95.1	2141	4	US-09-854-856-56
12	10211	94.4	2294	4	US-09-854-856-50
13	9929	91.8	1999	4	US-09-854-856-16
14	9872.5	91.3	2092	4	US-09-854-856-10
15	9796	90.6	2245	4	US-09-854-856-4
16	9765	90.3	1971	4	US-09-854-856-32
17	9708.5	89.8	2064	4	US-09-854-856-26
18	9632	89.1	2217	4	US-09-854-856-20
19	9627	89.0	1939	4	US-09-854-856-48
20	9570.5	88.5	2032	4	US-09-854-856-42
21	9494	87.8	2185	4	US-09-854-856-36
22	9463	87.5	1911	4	US-09-854-856-64
23	9406.5	87.0	2004	4	US-09-854-856-58
24	9330	86.3	2157	4	US-09-854-856-52
25	4053	37.5	982	4	US-09-854-856-6
26	4043	37.4	829	4	US-09-854-856-12
27	3889	35.0	954	4	US-09-854-856-22

28	3879	35.9	801	4	US-09-854-856-38	Sequence 28, Appl
29	3751	34.7	922	4	US-09-854-856-38	Sequence 28, Appl
30	3741	34.6	769	4	US-09-854-856-34	Sequence 44, Appl
31	3587	33.2	894	4	US-09-854-856-44	Sequence 54, Appl
32	3577	33.1	741	4	US-09-854-856-60	Sequence 60, Appl
33	1677	15.5	324	4	US-09-205-258-53	Sequence 53, Appl
34	1626	15.0	1601	4	US-09-342-473E-40	Sequence 40, Appl
35	1323	12.2	309	4	US-09-342-473E-37	Sequence 37, Appl
36	1280.5	11.8	1381	4	US-09-808-701A-25	Sequence 25, Appl
37	761.5	7.0	613	4	US-09-342-473E-39	Sequence 39, Appl
38	745.5	6.9	516	4	US-09-342-473E-34	Sequence 34, Appl
39	702.5	6.5	549	4	US-09-342-473E-32	Sequence 32, Appl
40	678.5	6.3	677	4	US-09-342-473E-38	Sequence 38, Appl
41	672	6.2	133	4	US-09-205-258-53	Sequence 53, Appl
42	649	6.0	138	1	US-08-664-596B-22	Sequence 22, Appl
43	439	4.1	2468	4	US-09-976-5594-726	Sequence 726, Appl
44	411	3.8	1239	4	US-09-688-188B-13	Sequence 13, Appl
45	411	3.8	1239	4	US-09-291-417D-13	Sequence 13, Appl

ALIGNMENTS

```

US-RESULT 1
US-09-854-856-14
: Sequence 14, Application US/09854856
: Patent No. 6541252
: GENERAL INFORMATION:
: APPLICANT: Walke, D. Wade
: APPLICANT: Hilbun, Erin
: APPLICANT: Donoho, Gregory
: APPLICANT: Turner, C. Alexander Jr.
: TITLE OF INVENTION: No. 6541252el Human Kinases and Polynucleotides
: TITLE OF INVENTION: Encoding the Same
: FILE REFERENCE: LEX-0178-USA
: CURRENT APPLICATION NUMBER: US/09/854,856
: CURRENT FILING DATE: 2001-05-14
: PRIOR APPLICATION NUMBER: US 60/206,015
: PRIOR FILING DATE: 2000-05-19
: NUMBER OF SEQ ID NOS: 64
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 14
: LENGTH: 2136
: TYPE: PRT
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: VARIANT
: LOCATION: (1)...(2136)
: OTHER INFORMATION: Xaa = Any Amino Acid
US-09-854-856-14

Query Match 100.0%; Score 10810; DB 4; Length 2136;
Best Local Similarity 100.0%; Pred. NO. 0;
Matches 2136; Conservative 0; Mismatches 0; Indels 0; Gaps 0

QY 1 MSGAAEKOSSTPSSLFLSPAPAPAKNGSSSDSSVGEKLGAAAADAVTGRTEEYRRRHT 60
Db 1 MSGGAEEKOSTPSSLFLSPAPAPAKNGSSSDSSVGEKLGAAAADAVTGRTEEYRRRHT 60
QY 61 MDKDSRGAAATTTTHERFFRRSVYCDSNATALELPGLPPLSPQPSITAAVPQSAPEPPH 120
Db 61 MDKDSRGAAATTTTHERFFRRSVYCDSNATALELPGLPPLSPQPSITAAVPQSAPEPPH 120
QY 121 REEYTTATATSOVAAQPPAAAPAGQAVAGAPSTVPSSTSKDRFVSQPSLVGSKKEEPP 180
Db 121 REEYTTATATSOVAAQPPAAAPAGQAVAGAPSTVPSSTSKDRFVSQPSLVGSKKEEPP 180
QY 121 REEYTTATATSOVAAQPPAAAPAGQAVAGAPSTVPSSTSKDRFVSQPSLVGSKKEEPP 180
Db 121 REEYTTATATSOVAAQPPAAAPAGQAVAGAPSTVPSSTSKDRFVSQPSLVGSKKEEPP 180
QY 181 ARSGGGGSAKEPQERSQQDDIELEETKAVGMSNDRFLKFDIEIGRGSFKTVYKGLD 240
Db 181 ARSGGGGSAKEPQERSQQDDIELEETKAVGMSNDRFLKFDIEIGRGSFKTVYKGLD 240
QY 241 TETTYEVAWCELDPRKLTKSERQPKKEAEMLKGLQHNNITRFYDWSWESTYKGKCTIVY 300
Db 241 TETTYEVAWCELDPRKLTKSERQPKKEAEMLKGLQHNNITRFYDWSWESTYKGKCTIVY 300

```

Db 241 TETVEVAMCELODRKLTSEORFKEBAEMLKGLHFNIVRFVDSWESVTKKCIYLV 300
 QY 301 TELMTSSGLKTYLTKRFKXMKTKVLBSNCRQILKGLQFHTRPPIIHDLCNDIETFGP 360
 Db 301 TELMTSSGLKTYLTKRFKXMKTKVLBSNCRQILKGLQFHTRPPIIHDLCNDIETFGP 360
 QY 361 TGSVKGIDGLATLKRASFVKSIGTPEPFMAPEMEYEKYDESVDYAFGCMLEVAATSEY 420
 Db 361 TGSVKGIDGLATLKRASFVKSIGTPEPFMAPEMEYEKYDESVDYAFGCMLEVAATSEY 420
 QY 421 PYSEQNAAOIYRRVTSVGVKPAEDKVAIPEVKEIIEGCIKONKDERYSIKDLNHAFFQ 480
 Db 421 PYSEQNAAOIYRRVTSVGVKPAEDKVAIPEVKEIIEGCIKONKDERYSIKDLNHAFFQ 480
 QY 481 BETGVRALEAEDDOEKALIKMLRIEDIKLKGYKXNEALIEPSFDLERVPEVNAQEM 540
 Db 481 BETGVRALEAEDDOEKALIKMLRIEDIKLKGYKXNEALIEPSFDLERVPEVNAQEM 540
 QY 541 VESGVCEGDHKTAKAIKDRVSLIKRKEQROLVREOEKKKQESSLKQOVESSASQ 600
 Db 541 VESGVCEGDHKTAKAIKDRVSLIKRKEQROLVREOEKKKQESSLKQOVESSASQ 600
 QY 601 TGTIQLBSASTGICTASTTSASVSTQVEPEPEADQHQOLQYQOPISVLSGTVDSGQ 660
 Db 601 TGTIQLBSASTGICTASTTSASVSTQVEPEPEADQHQOLQYQOPISVLSGTVDSGQ 660
 QY 661 SSVFTESVSSQOQVSVGSQHEOAHSTGVPGHITSTTQAOQSPHGVYPPSSVAGQSGSQ 720
 Db 661 SSVFTESVSSQOQVSVGSQHEOAHSTGVPGHITSTTQAOQSPHGVYPPSSVAGQSGSQ 720
 QY 721 QPSSSLTGVSSSQPIQHPOQOOGIIOQTAPPOQTQVXSLSQTSSEATTAQVPSOPQAP 780
 Db 721 QPSSSLTGVSSSQPIQHPOQOOGIIOQTAPPOQTQVXSLSQTSSEATTAQVPSOPQAP 780
 QY 781 QVLPQVAGAKSITGVSGVAPAEFVAVAQPOATPPTLASSVDSAHSDVAGMSDGENV 840
 Db 781 QVLPQVAGAKSITGVSGVAPAEFVAVAQPOATPPTLASSVDSAHSDVAGMSDGENV 840
 QY 841 PSSSGRIEGRTTKHYKSVRSRHEKTSRPKRIINVSNGRVRVECOLETNRKMYT 900
 Db 841 PSSSGRIEGRTTKHYKSVRSRHEKTSRPKRIINVSNGRVRVECOLETNRKMYT 900
 QY 901 FKFDLDGDNPEELATIMVNDPILAIERESFVDQVREIIEKADMLSEDSVEBEGQGL 960
 Db 901 FKFDLDGDNPEELATIMVNDPILAIERESFVDQVREIIEKADMLSEDSVEBEGQGL 960
 QY 961 ESLOGKDDYGFSGQKLEGEFKQPIPASMPQOIGITSSITQVHAGRRFIVSPVES 1020
 Db 961 ESLOGKDDYGFSGQKLEGEFKQPIPASMPQOIGITSSITQVHAGRRFIVSPVES 1020
 QY 1021 RLRESKVFSEITDTTVAASTAOSGKMLSHSASSLSIQAFSELRRKQMBEGPTAPNF 1080
 Db 1021 RLRESKVFSEITDTTVAASTAOSGKMLSHSASSLSIQAFSELRRKQMBEGPTAPNF 1080
 QY 1081 SHTGPIFPVPPPLSSIAGVPTTAATAATAPVATSSPNDISTSVTQSEVTPTEGIAGV 1140
 Db 1081 SHTGPIFPVPPPLSSIAGVPTTAATAATAPVATSSPNDISTSVTQSEVTPTEGIAGV 1140
 QY 1141 ATSTGVTSGLPPIPVSESPVLSVSVSITTPAVNISITTSBGLVPTSTSEIVSSIA 1200
 Db 1141 ATSTGVTSGLPPIPVSESPVLSVSVSITTPAVNISITTSBGLVPTSTSEIVSSIA 1200
 QY 1201 LYSVTSVATSASAGSTATPGPKPVPVVSQAAGSTTVGATLTSVTTSPSTASOLS 1260
 Db 1201 LYSVTSVATSASAGSTATPGPKPVPVVSQAAGSTTVGATLTSVTTSPSTASOLS 1260
 QY 1261 IQSSSSTSTPTLAETVVASHSLDKTSHSSTTGLAFSLAPSSSSSPGAGVSYISQPG 1320
 Db 1261 IQSSSSTSTPTLAETVVASHSLDKTSHSSTTGLAFSLAPSSSSSPGAGVSYISQPG 1320
 QY 1321 LHLPLVPSVASTPILPOAAGPTSTPLIPVPSIPPLVQPVANVPAVOQTLHSQOPAL 1380
 Db 1321 LHLPLVPSVASTPILPOAAGPTSTPLIPVPSIPPLVQPVANVPAVOQTLHSQOPAL 1380

QY 1381 LPNPHHCPEVSDTOPKAPGIDDIKTLKEKLSLSESHSSGAQHASVLETSLVIES 1440
 Db 1381 LPNPHHCPEVSDTOPKAPGIDDIKTLKEKLSLSESHSSGAQHASVLETSLVIES 1440
 QY 1441 TVTGPITTAAPSKILATSTSTCLPPTNLPLGVAAIPVTVPVPGOVSTVSTTSVK 1500
 Db 1441 TVTGPITTAAPSKILATSTSTCLPPTNLPLGVAAIPVTVPVPGOVSTVSTTSVK 1500
 QY 1501 PGTAPEPLTKAVLPVTELPAGTLPESEQLPPEPGSILTSOQOPLDLDLAOLRTTSP 1560
 Db 1501 PGTAPEPLTKAVLPVTELPAGTLPESEQLPPEPGSILTSOQOPLDLDLAOLRTTSP 1560
 QY 1561 EXITVTGAVGVSAAPALITAEAGTOQKGVSVYKEGVALTSSGAGYFKKGRPOVSAA 1620
 Db 1561 EXITVTGAVGVSAAPALITAEAGTOQKGVSVYKEGVALTSSGAGYFKKGRPOVSAA 1620
 QY 1621 DGAQKEGKNKSEDAKSVFSESTSESSVLSSSSPSTLVKPEPNGITTPGISDVPESAH 1680
 Db 1621 DGAQKEGKNKSEDAKSVFSESTSESSVLSSSSPSTLVKPEPNGITTPGISDVPESAH 1680
 QY 1681 KTTASAKSDTQPTKYGRFOVTTTANKVGRFSVSKTEDKITDTKKEGVPVAPPEFMLEQ 1740
 Db 1681 KTTASAKSDTQPTKYGRFOVTTTANKVGRFSVSKTEDKITDTKKEGVPVAPPEFMLEQ 1740
 QY 1741 AVLPAAVTPKKEKPELSEPSHNLNPSDPEAFLSRDVEDGSGSPHSQSLPSQNL 1800
 Db 1741 AVLPAAVTPKKEKPELSEPSHNLNPSDPEAFLSRDVEDGSGSPHSQSLPSQNL 1800
 QY 1801 SOSLSNFSNYSMSDNESDIEDEDLKLRLRLDKHKEIODLSRQKHEIESLYTKLG 1860
 Db 1801 SOSLSNFSNYSMSDNESDIEDEDLKLRLRLDKHKEIODLSRQKHEIESLYTKLG 1860
 QY 1861 KVPAAVILPPAALSGRRRPTKSGKSSRSLSGKNSPOLSGNLSGQSAASVLAHQOT 1920
 Db 1861 KVPAAVILPPAALSGRRRPTKSGKSSRSLSGKNSPOLSGNLSGQSAASVLAHQOT 1920
 QY 1921 LHPGNIPEEGQOOLQPLKPSBSDMLYSAFTSDGAISSPISAPQGTSTTVGATV 1980
 Db 1921 LHPGNIPEEGQOOLQPLKPSBSDMLYSAFTSDGAISSPISAPQGTSTTVGATV 1980
 QY 1981 NSQAQAQAPAMTSSRGFTTDLHLKLVDMWARDAMLSGRGSGKGMNTEGPMARKFS 2040
 Db 1981 NSQAQAQAPAMTSSRGFTTDLHLKLVDMWARDAMLSGRGSGKGMNTEGPMARKFS 2040
 QY 2041 APQOLCISMTSNLGSAPISAAATSLGHFTKSMCPQOYGFPATPGAGQMSGTGPAPQ 2100
 Db 2041 APQOLCISMTSNLGSAPISAAATSLGHFTKSMCPQOYGFPATPGAGQMSGTGPAPQ 2100
 QY 2101 PLGQPOVGTASLONFRISLQKISINPESNIRTT 2136
 Db 2101 PLGQPOVGTASLONFRISLQKISINPESNIRTT 2136

RESULT 2
 US-09-854-856-8
 ; Sequence 8, Application US/09854856
 ; Patent No. 6541252
 ; GENERAL INFORMATION:
 ; APPLICANT: Walke, D. Wade
 ; APPLICANT: Hilbun, Erin
 ; APPLICANT: Donoho, Gregory
 ; APPLICANT: Turner, C. Alexander Jr.
 ; TITLE OF INVENTION: No. 6541252 Human Kinases and Polynucleotides
 ; TITLE OF INVENTION: Encoding the Same
 ; FILE REFERENCE: Lex-0178-USA
 ; CURRENT APPLICATION NUMBER: US/09/854,856
 ; PRIOR APPLICATION NUMBER: US 60/206,015
 ; PRIOR FILING DATE: 2000-05-19
 ; NUMBER OF SEQ ID NOS: 64
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 8

LENGTH: 2229
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: VARIANT
LOCATION: (1)...(2229)
OTHER INFORMATION: Xaa = Any Amino Acid
US-09-854-856-8

Query Match 99.5%; Score 10753.5; DB 4; Length 2229;
Best Local Similarity 95.8%; Pred. No. 0;
Matches 2136; Conservative 0; Mismatches 0; Indels 93; Gaps 1;

1 MSGGAKEKOSTPSTSLFLSPAPAPAKNGSSSDSVGKLGAAADAVTGRTEYRRRRHT 60
1 MSGGAKEKOSTPSTSLFLSPAPAPAKNGSSSDSVGKLGAAADAVTGRTEYRRRRHT 60
61 MDKSRGAAATTTTTHRRFRRSVICSNAITALEPGLPLSLPQPSIPAAVPGAPPEPH 120
61 MDKSRGAAATTTTTHRRFRRSVICSNAITALEPGLPLSLPQPSIPAAVPGAPPEPH 120
121 REETVATATSOVAQOQPPAAAPGEQAVAGPAPSTVSSSTSKDPVQPSLVSKKEPP 180
121 REETVATATSOVAQOQPPAAAPGEQAVAGPAPSTVSSSTSKDPVQPSLVSKKEPP 180
181 ARSGGGGSAKEPQEESSQODDIELETAKVGMNDGRFLKFDIEIGRSGFKTVYKGLD 240
181 ARSGGGGSAKEPQEESSQODDIELETAKVGMNDGRFLKFDIEIGRSGFKTVYKGLD 240
241 TETTVAVAMCELQDRKLTSEKORFKKEAEMLKGLQHPNIVRFYDSWSTVKGKCIYLV 300
241 TETTVAVAMCELQDRKLTSEKORFKKEAEMLKGLQHPNIVRFYDSWSTVKGKCIYLV 300
301 TELMTSGTLKTYLKRFKYMKIYVRSNCROILKGIQPLHTPTPIIRDLKCDNIITGP 360
301 TELMTSGTLKTYLKRFKYMKIYVRSNCROILKGIQPLHTPTPIIRDLKCDNIITGP 360
361 TGSYKIGLGLATIKRASFAKSVIGTFPMAPEMEYEEKYDSVVVYAFGCMLEMAISEY 420
361 TGSYKIGLGLATIKRASFAKSVIGTFPMAPEMEYEEKYDSVVVYAFGCMLEMAISEY 420
421 PYSECQNAQAQYRRVTSVGPAPSPDKVAIPVEKELIEGCIKONDERYSIKDLINHAFFQ 480
421 PYSECQNAQAQYRRVTSVGPAPSPDKVAIPVEKELIEGCIKONDERYSIKDLINHAFFQ 480
481 BETGVAVLAEDDGEKTAIKMLRIEDIKLKGKYNKNEALIESFDLERVPEDEVAGEM 540
481 BETGVAVLAEDDGEKTAIKMLRIEDIKLKGKYNKNEALIESFDLERVPEDEVAGEM 540
541 VESGVGCEGDHKTAKAIKDRVSLIKRKREORQVREBQKKKQESSLKOQVESSASQ 600
541 VESGVGCEGDHKTAKAIKDRVSLIKRKREORQVREBQKKKQESSLKOQVESSASQ 600
601 TGKQLPASTGIPASTTSASVSTQVEPEEADQOQLOQOQPSIVLSDGTVDGSGG 660
601 TGKQLPASTGIPASTTSASVSTQVEPEEADQOQLOQOQPSIVLSDGTVDGSGG 660
661 SSVFPESSVSSQQTYSYSGHQAHSSTGTVPCHIPSTVQAOQOPRGVPPSSVAGSGSG 720
661 SSVFPESSVSSQQTYSYSGHQAHSSTGTVPCHIPSTVQAOQOPRGVPPSSVAGSGSG 720
721 QPSSSLTGVSSSQIPIHPQOQOQIQOTAPPOQTQYSLSQTSTSEATTAQPVQAPAP 780
721 QPSSSLTGVSSSQIPIHPQOQOQIQOTAPPOQTQYSLSQTSTSEATTAQPVQAPAP 780
781 QVLPOVSAQKQ----- 791
781 QVLPOVSAQKQ----- 791
791 QVLPOVSAQKQ----- 791
791 QVLPOVSAQKQ----- 791
841 VVQPVESNLIVPMGVGGQVQVQSPGGSIAQAPTTSSQQAVALSTQGVSAVAPAEVAV 900
841 VVQPVESNLIVPMGVGGQVQVQSPGGSIAQAPTTSSQQAVALSTQGVSAVAPAEVAV 900

808 ACPQATOPTTLASSVDASDVASGMSDGNENPVSSSGRRHGRTRKRYKKSVRSRHRE 867
808 ACPQATOPTTLASSVDASDVASGMSDGNENPVSSSGRRHGRTRKRYKKSVRSRHRE 867
901 ACPQATOPTTLASSVDASDVASGMSDGNENPVSSSGRRHGRTRKRYKKSVRSRHRE 960
901 ACPQATOPTTLASSVDASDVASGMSDGNENPVSSSGRRHGRTRKRYKKSVRSRHRE 960
868 KTSRPKRLITLVNKNKDRVVECOLETHNRKMYTFKFDLODNDPEELATIVNDFILAE 927
868 KTSRPKRLITLVNKNKDRVVECOLETHNRKMYTFKFDLODNDPEELATIVNDFILAE 927
961 KTSRPKRLITLVNKNKDRVVECOLETHNRKMYTFKFDLODNDPEELATIVNDFILAE 1020
961 KTSRPKRLITLVNKNKDRVVECOLETHNRKMYTFKFDLODNDPEELATIVNDFILAE 1020
928 RESFVQVREIIEKADMELESDVSEPEEGDQLESIGKQDXYFGSGOKLEGEKOPIPA 987
928 RESFVQVREIIEKADMELESDVSEPEEGDQLESIGKQDXYFGSGOKLEGEKOPIPA 987
1021 RESFVQVREIIEKADMELESDVSEPEEGDQLESIGKQDXYFGSGOKLEGEKOPIPA 1080
1021 RESFVQVREIIEKADMELESDVSEPEEGDQLESIGKQDXYFGSGOKLEGEKOPIPA 1080
988 GSMPOQIGIPTSSLTVGVHAGRRFIVSPVESRLBRSKVPSEBITDTVAASTAQSQGM 1140
988 GSMPOQIGIPTSSLTVGVHAGRRFIVSPVESRLBRSKVPSEBITDTVAASTAQSQGM 1140
1081 GSMPOQIGIPTSSLTVGVHAGRRFIVSPVESRLBRSKVPSEBITDTVAASTAQSQGM 1140
1081 GSMPOQIGIPTSSLTVGVHAGRRFIVSPVESRLBRSKVPSEBITDTVAASTAQSQGM 1140
1048 LSHSASLSIQOAFSELRAROMTEGPTAPNNSHGPTEPPVPPFLSSJAGVTTAAAT 1107
1048 LSHSASLSIQOAFSELRAROMTEGPTAPNNSHGPTEPPVPPFLSSJAGVTTAAAT 1107
1141 LSHSASLSIQOAFSELRAROMTEGPTAPNNSHGPTEPPVPPFLSSJAGVTTAAAT 1200
1141 LSHSASLSIQOAFSELRAROMTEGPTAPNNSHGPTEPPVPPFLSSJAGVTTAAAT 1200
1108 APVPATSSPNDISTSVIQSEVTPTEEGIAGVATSTGVVTSGLPIIPVSESPVLSV 1167
1108 APVPATSSPNDISTSVIQSEVTPTEEGIAGVATSTGVVTSGLPIIPVSESPVLSV 1167
1201 APVPATSSPNDISTSVIQSEVTPTEEGIAGVATSTGVVTSGLPIIPVSESPVLSV 1260
1201 APVPATSSPNDISTSVIQSEVTPTEEGIAGVATSTGVVTSGLPIIPVSESPVLSV 1260
1168 SSITTPAVASISTTSPELQVPTSTSEIVVSTALYPSVYVATSASAGSSTATPGKPPA 1227
1168 SSITTPAVASISTTSPELQVPTSTSEIVVSTALYPSVYVATSASAGSSTATPGKPPA 1227
1261 SSITTPAVASISTTSPELQVPTSTSEIVVSTALYPSVYVATSASAGSSTATPGKPPA 1320
1261 SSITTPAVASISTTSPELQVPTSTSEIVVSTALYPSVYVATSASAGSSTATPGKPPA 1320
1228 VWSQQAAGSTTVGATLTVSSTTSPSTASOLSTQSSSTPPLATVWVSAHSLKTS 1287
1228 VWSQQAAGSTTVGATLTVSSTTSPSTASOLSTQSSSTPPLATVWVSAHSLKTS 1287
1321 VWSQQAAGSTTVGATLTVSSTTSPSTASOLSTQSSSTPPLATVWVSAHSLKTS 1380
1321 VWSQQAAGSTTVGATLTVSSTTSPSTASOLSTQSSSTPPLATVWVSAHSLKTS 1380
1288 HSSSTGLAFSLAPSSSSPGAGVSYISOPGHLPIVSVIATPILPQAGPTSTPL 1347
1288 HSSSTGLAFSLAPSSSSPGAGVSYISOPGHLPIVSVIATPILPQAGPTSTPL 1347
1381 HSSSTGLAFSLAPSSSSPGAGVSYISOPGHLPIVSVIATPILPQAGPTSTPL 1440
1381 HSSSTGLAFSLAPSSSSPGAGVSYISOPGHLPIVSVIATPILPQAGPTSTPL 1440
1348 LPQVPSIPPLVQPVANPAVQOQTLHSQPOPALLPNOPTHCEVDSDTQPKAGIDIX 1407
1348 LPQVPSIPPLVQPVANPAVQOQTLHSQPOPALLPNOPTHCEVDSDTQPKAGIDIX 1407
1441 LPQVPSIPPLVQPVANPAVQOQTLHSQPOPALLPNOPTHCEVDSDTQPKAGIDIX 1500
1441 LPQVPSIPPLVQPVANPAVQOQTLHSQPOPALLPNOPTHCEVDSDTQPKAGIDIX 1500
1408 TLEKRLSLPSHSSSGAGHASVLETSIVTSGIPPTAAVAPSKLSTSTGCLP 1467
1408 TLEKRLSLPSHSSSGAGHASVLETSIVTSGIPPTAAVAPSKLSTSTGCLP 1467
1501 TLEKRLSLPSHSSSGAGHASVLETSIVTSGIPPTAAVAPSKLSTSTGCLP 1560
1501 TLEKRLSLPSHSSSGAGHASVLETSIVTSGIPPTAAVAPSKLSTSTGCLP 1560
1468 TNLPLGTVALPVTVPVTPGQVSTPSTTSVSKFGTAISKXPLKAPVLPVGTBLPAGTL 1527
1468 TNLPLGTVALPVTVPVTPGQVSTPSTTSVSKFGTAISKXPLKAPVLPVGTBLPAGTL 1527
1561 TNLPLGTVALPVTVPVTPGQVSTPSTTSVSKFGTAISKXPLKAPVLPVGTBLPAGTL 1620
1561 TNLPLGTVALPVTVPVTPGQVSTPSTTSVSKFGTAISKXPLKAPVLPVGTBLPAGTL 1620
1528 PSEQLPPEPGPSLTQSOQPLEDDLAQRLRTLSPEXITVTSVAVGVSAAPALITBAGTQ 1587
1528 PSEQLPPEPGPSLTQSOQPLEDDLAQRLRTLSPEXITVTSVAVGVSAAPALITBAGTQ 1587
1621 PSEQLPPEPGPSLTQSOQPLEDDLAQRLRTLSPEXITVTSVAVGVSAAPALITBAGTQ 1680
1621 PSEQLPPEPGPSLTQSOQPLEDDLAQRLRTLSPEXITVTSVAVGVSAAPALITBAGTQ 1680
1588 QKGVSOVKEGPIVLAATSSGAGVFKMGFRQVSAADGAQKEGKNSKEDASKVFESESTSS 1647
1588 QKGVSOVKEGPIVLAATSSGAGVFKMGFRQVSAADGAQKEGKNSKEDASKVFESESTSS 1647
1681 QKGVSOVKEGPIVLAATSSGAGVFKMGFRQVSAADGAQKEGKNSKEDASKVFESESTSS 1740
1681 QKGVSOVKEGPIVLAATSSGAGVFKMGFRQVSAADGAQKEGKNSKEDASKVFESESTSS 1740
1648 VLSSSPESTLVKEPENGITIPGISDVPEBAHKTITASBAKSDTGOPTKVRFOVTTTAN 1707
1648 VLSSSPESTLVKEPENGITIPGISDVPEBAHKTITASBAKSDTGOPTKVRFOVTTTAN 1707
1741 VLSSSPESTLVKEPENGITIPGISDVPEBAHKTITASBAKSDTGOPTKVRFOVTTTAN 1800
1741 VLSSSPESTLVKEPENGITIPGISDVPEBAHKTITASBAKSDTGOPTKVRFOVTTTAN 1800
1708 KVGRRSVSKTEDEKTDITKKEGPVAPPPMDLEQAVLPVAVIKKEPELSRFSHNLGPPSD 1767
1708 KVGRRSVSKTEDEKTDITKKEGPVAPPPMDLEQAVLPVAVIKKEPELSRFSHNLGPPSD 1767
1801 KVGRRSVSKTEDEKTDITKKEGPVAPPPMDLEQAVLPVAVIKKEPELSRFSHNLGPPSD 1860
1801 KVGRRSVSKTEDEKTDITKKEGPVAPPPMDLEQAVLPVAVIKKEPELSRFSHNLGPPSD 1860
1768 PEAFLASDVDDGSGSPHSPIHLSKSLPSQNLISQLSNSFNSSVMSDNDSDIEDEDLX 1827
1768 PEAFLASDVDDGSGSPHSPIHLSKSLPSQNLISQLSNSFNSSVMSDNDSDIEDEDLX 1827
1861 PEAFLASDVDDGSGSPHSPIHLSKSLPSQNLISQLSNSFNSSVMSDNDSDIEDEDLX 1920
1861 PEAFLASDVDDGSGSPHSPIHLSKSLPSQNLISQLSNSFNSSVMSDNDSDIEDEDLX 1920
1828 LELRLRLDKHLKEIQLDLSRQKHETESLYTKLVKVPVAVIIPPAAPLGGRRRRPTKSGS 1887
1828 LELRLRLDKHLKEIQLDLSRQKHETESLYTKLVKVPVAVIIPPAAPLGGRRRRPTKSGS 1887
1921 LELRLRLDKHLKEIQLDLSRQKHETESLYTKLVKVPVAVIIPPAAPLGGRRRRPTKSGS 1980
1921 LELRLRLDKHLKEIQLDLSRQKHETESLYTKLVKVPVAVIIPPAAPLGGRRRRPTKSGS 1980
1888 KSSRSSSLGNKSPOLSGNLSCQSAASVLAHPQOQTLHPNGNIPESGQNLQPLKPSBSN 1947
1888 KSSRSSSLGNKSPOLSGNLSCQSAASVLAHPQOQTLHPNGNIPESGQNLQPLKPSBSN 1947

```
Db 1981 KSSRSSISGKNSPQJSGNSISGQSAASVLPQCTLHPGNIPESSGQNLQPLKPSPSN 2040
QY 1948 LYSATSPGALSVPLSAPGQCTSSNTVGAIVNSQAQAOPAPMTSSKGTFTDHLK 2007
Db 2041 LYSATSPGALSVPLSAPGQCTSSNTVGAIVNSQAQAOPAPMTSSKGTFTDHLK 2100
QY 2008 VDNMARDAANLSGRGSKGNNYEGPMARKFSAPGOLCISMTSNLGGSAPIASAATSL 2067
Db 2101 VDNMARDAANLSGRGSKGNNYEGPMARKFSAPGOLCISMTSNLGGSAPIASAATSL 2160
QY 2068 GHFTSMCPPOQYGPATPFGAQMSGTGAPAPQPGQPVGTASLQNFNLSLQKISLN 2127
Db 2161 GHFTSMCPPOQYGPATPFGAQMSGTGAPAPQPGQPVGTASLQNFNLSLQKISLN 2220
QY 2128 PPGSNLRAT 2136
Db 2221 PPGSNLRAT 2229

RESULT 3
US-09-854-856-2
; Sequence 2, Application US/09854856
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Hilbun, Erin
; APPLICANT: Donoho, Gregory
; APPLICANT: Turner, C. Alexander Jr.
; TITLE OF INVENTION: No. 6541252el Human Kinases and Polynucleotides
; FILE REFERENCE: LEX-0178-USA
; CURRENT FILING DATE: 2001-05-14
; PRIOR FILING DATE: 2000-05-19
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 2382
; TYPE: PR
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(2382)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-09-854-856-2

Query Match 96.8%; Score 10677; DB 4; Length 2382;
Best Local Similarity 89.7%; Pred. No. 0;
Matches 2136; Conservative 0; Mismatches 0; Indels 246; Gaps 1;

QY 1 MSGGAAGKOSTPSTPLSPAPAPKNGSSSDSVGKLGAAADANTGTEERRRRHT 60
Db 1 MSGGAAGKOSTPSTPLSPAPAPKNGSSSDSVGKLGAAADANTGTEERRRRHT 60
QY 61 MDXDSRGAATTTTTEHFRFRSVICDSNATALELPGPLSLPSPSIPAAVPOSAPREP 120
Db 61 MDXDSRGAATTTTTEHFRFRSVICDSNATALELPGPLSLPSPSIPAAVPOSAPREP 120
QY 121 REETVNTATSOVAQAOPPAAPAAPEQAIVAGPAPSTVSSKDPVPSQPSLVGSKKEPP 180
Db 121 REETVNTATSOVAQAOPPAAPAAPEQAIVAGPAPSTVSSKDPVPSQPSLVGSKKEPP 180
QY 181 ARSGSGGSAKEPQERSQOQDDELEETKAVGMSNGRFLKPDIEGRSSFTVYKGLD 240
Db 181 ARSGSGGSAKEPQERSQOQDDELEETKAVGMSNGRFLKPDIEGRSSFTVYKGLD 240
QY 241 TETTVAVMCELQDKLTSEKRFKEEAMLGLOHPNIVRFYDSWESTVKKKCIIV 300
Db 241 TETTVAVMCELQDKLTSEKRFKEEAMLGLOHPNIVRFYDSWESTVKKKCIIV 300
QY 301 TELMTSGTLTKYKRFKVMKIKVLRSCWQIQLGLOFLHTRTPPIIHRDLKCNIFITGP 360
```

```
Db 301 TELMTSGTLTKYKRFKVMKIKVLRSCWQIQLGLOFLHTRTPPIIHRDLKCNIFITGP 360
QY 361 TGSVYKIGLGLATLKRASFANSVIGTPEFMAPEMEYEKYDESVDYAFGCMLEWATSEY 420
Db 361 TGSVYKIGLGLATLKRASFANSVIGTPEFMAPEMEYEKYDESVDYAFGCMLEWATSEY 420
QY 421 PYSECQNAQIYRYRYSVGVKPAKPAKPAKPAKPAKPAKPAKPAKPAKPAKPAKPAK 480
Db 421 PYSECQNAQIYRYRYSVGVKPAKPAKPAKPAKPAKPAKPAKPAKPAKPAKPAKPAK 480
QY 481 EETGVYVLAEDDEGEKAIKMLRIEDIKTKGKYKNEALFEFPLERDPEDVADQM 540
Db 481 EETGVYVLAEDDEGEKAIKMLRIEDIKTKGKYKNEALFEFPLERDPEDVADQM 540
QY 541 VESGYVCEGDHKTMAKAIKDRVSLIKRREOROLVREEQEKKKQESSLKOQVBOSSAQ 600
Db 541 VESGYVCEGDHKTMAKAIKDRVSLIKRREOROLVREEQEKKKQESSLKOQVBOSSAQ 600
QY 601 TGIKQLPASTGIPATSTASVSTQVEPEPEADQHQOLOYQOPSLISLSDGTVDGSG 660
Db 601 TGIKQLPASTGIPATSTASVSTQVEPEPEADQHQOLOYQOPSLISLSDGTVDGSG 660
QY 661 SSVFTESSVSOQTVSYSGOHEQASTGVPGHLPSTYQAOQPHGVYPPSSVAQOQSG 720
Db 661 SSVFTESSVSOQTVSYSGOHEQASTGVPGHLPSTYQAOQPHGVYPPSSVAQOQSG 720
QY 721 QPSSSLTVGVSSQPIQHPQOQOQIQTAPPOQTVQVSLQSTSTSEATTQAQVSPQAP 780
Db 721 QPSSSLTVGVSSQPIQHPQOQOQIQTAPPOQTVQVSLQSTSTSEATTQAQVSPQAP 780
QY 781 QVLPQVSAQKQ----- 791
Db 781 QVLPQVSAQKQPVSPVPTTQGPOLPVAQPSVPHSGHFLPVGQPLPTPLPQYP 840
QY 792 ----- 791
Db 792 ----- 791
QY 841 VSOPISTPHVSTAQTGFSPLITMAGITPPLTLTASSATTAAIPGVSTVPSQLPTLL 900
Db 841 VSOPISTPHVSTAQTGFSPLITMAGITPPLTLTASSATTAAIPGVSTVPSQLPTLL 900
QY 901 QPVTLPSQVHPQLQPAVQSMGIPANLGOAAEVLSSGDLVYGGFPPPLPQYRPSNI 960
Db 901 QPVTLPSQVHPQLQPAVQSMGIPANLGOAAEVLSSGDLVYGGFPPPLPQYRPSNI 960
QY 992 ----- 791
Db 992 ----- 791
QY 961 ABSNNVASCISHSTVLAPEMPEVLAATPGYFPYVQPVESNLLVPKGVGGVQVVSQPG 1020
Db 961 ABSNNVASCISHSTVLAPEMPEVLAATPGYFPYVQPVESNLLVPKGVGGVQVVSQPG 1020
QY 792 -----STQVSGVAPABPVAVVAPQATOPTTLASSVDSASHSVASGMS 834
Db 792 -----STQVSGVAPABPVAVVAPQATOPTTLASSVDSASHSVASGMS 834
QY 1021 GSILAQAPTTSSQQAVALSTQGVQVAPAEVVAQAOPQATPTTLASSVDSASHSVASGMS 1080
Db 1021 GSILAQAPTTSSQQAVALSTQGVQVAPAEVVAQAOPQATPTTLASSVDSASHSVASGMS 1080
QY 835 DGNENVPSSSGRHGRTTKKHYKYSRSRHEKTSBKPLILNVSNKGRVVECOLETH 894
Db 835 DGNENVPSSSGRHGRTTKKHYKYSRSRHEKTSBKPLILNVSNKGRVVECOLETH 894
QY 1081 DGNENVPSSSGRHGRTTKKHYKYSRSRHEKTSBKPLILNVSNKGRVVECOLETH 1140
Db 1081 DGNENVPSSSGRHGRTTKKHYKYSRSRHEKTSBKPLILNVSNKGRVVECOLETH 1140
QY 895 NRKMTVEKPLDDNBEIATITVNNDFILAIRESVVDVREIIEKADMLSEBVSVEP 954
Db 895 NRKMTVEKPLDDNBEIATITVNNDFILAIRESVVDVREIIEKADMLSEBVSVEP 954
QY 1141 NRKMTVEKPLDDNBEIATITVNNDFILAIRESVVDVREIIEKADMLSEBVSVEP 1200
Db 1141 NRKMTVEKPLDDNBEIATITVNNDFILAIRESVVDVREIIEKADMLSEBVSVEP 1200
QY 955 EGDQGLSLOGKNDYGRSGQKLEGEFKQPIPASSMPQOIGITPSSLTQVYHAGRRFTV 1014
Db 955 EGDQGLSLOGKNDYGRSGQKLEGEFKQPIPASSMPQOIGITPSSLTQVYHAGRRFTV 1014
QY 1201 EGDQGLSLOGKNDYGRSGQKLEGEFKQPIPASSMPQOIGITPSSLTQVYHAGRRFTV 1260
Db 1201 EGDQGLSLOGKNDYGRSGQKLEGEFKQPIPASSMPQOIGITPSSLTQVYHAGRRFTV 1260
QY 1015 SPVPESRLRESKYFPESEITDTVAASTAOSFGNMLSHSASSLSQOAFSELRAAQMTGPN 1074
Db 1015 SPVPESRLRESKYFPESEITDTVAASTAOSFGNMLSHSASSLSQOAFSELRAAQMTGPN 1074
QY 1261 SPVPESRLRESKYFPESEITDTVAASTAOSFGNMLSHSASSLSQOAFSELRAAQMTGPN 1320
Db 1261 SPVPESRLRESKYFPESEITDTVAASTAOSFGNMLSHSASSLSQOAFSELRAAQMTGPN 1320
QY 1075 TAPNFSHGTPFPVVPPLFSSITAGVTTAAATAPVATSSPNDISTVYQSEVVPTE 1134
Db 1075 TAPNFSHGTPFPVVPPLFSSITAGVTTAAATAPVATSSPNDISTVYQSEVVPTE 1134
QY 1321 TAPNFSHGTPFPVVPPLFSSITAGVTTAAATAPVATSSPNDISTVYQSEVVPTE 1380
Db 1321 TAPNFSHGTPFPVVPPLFSSITAGVTTAAATAPVATSSPNDISTVYQSEVVPTE 1380
QY 1135 EGIAGVASTGVVTSGLPIPVSESPVLSVVSSTITPAVVSISTTSGLQVPTSTSEI 1194
Db 1135 EGIAGVASTGVVTSGLPIPVSESPVLSVVSSTITPAVVSISTTSGLQVPTSTSEI 1194
```

Db 1381 EGIAGVATSGVTSGGLPIPVSESEFVLSVSSITIPAVVISITSPSLQVETSTSEI 1440
QY 1195 VVSSTALYPSVTYATSSASAGSSTATPGPKPPAVVSOOAGSTTVGATLTSTSTTSPS 1254
Db 1441 VVSSTALYPSVTYATSSASAGSSTATPGPKPPAVVSOOAGSTTVGATLTSTSTTSPS 1500
QY 1255 TAAQLSTIQLSSSTSTPTLAETVVVSAHSLDKTSHSSTTGAFLSAPSSSSSPGAGVSSY 1314
Db 1501 TAAQLSTIQLSSSTSTPTLAETVVVSAHSLDKTSHSSTTGAFLSAPSSSSSPGAGVSSY 1560
QY 1315 ISQGGHPLVYBESVIASTPLLPQAAGPTSTPLLPQVPSIPPLVQPVANPAVQOQLIHS 1374
Db 1561 ISQGGHPLVYBESVIASTPLLPQAAGPTSTPLLPQVPSIPPLVQPVANPAVQOQLIHS 1620
QY 1375 QPOPALPNOPHTCPEVSDTOPKAGIDIDIKLEELKRLSEHSSSGAQAHSVLEET 1434
Db 1621 QPOPALPNOPHTCPEVSDTOPKAGIDIDIKLEELKRLSEHSSSGAQAHSVLEET 1680
QY 1435 SLVIESITVTPGIPPTTAAVAPSKLTSTSTCLPPTNLPGLVALPVTPVTPGQVSTPVST 1494
Db 1681 SLVIESITVTPGIPPTTAAVAPSKLTSTSTCLPPTNLPGLVALPVTPVTPGQVSTPVST 1740
QY 1495 TTSQVKEGTAPSKPKPLTKAPVLPVGTBLPAGTLPEEQLPPEPGLTQSQQLPDLDAQL 1554
Db 1741 TTSQVKEGTAPSKPKPLTKAPVLPVGTBLPAGTLPEEQLPPEPGLTQSQQLPDLDAQL 1800
QY 1555 RRTISPEKITVTSVAVGPVMAAPTAITEAGTQPOKGVQVKEGPVLATSSAGVFKMGKF 1614
Db 1801 RRTISPEKITVTSVAVGPVMAAPTAITEAGTQPOKGVQVKEGPVLATSSAGVFKMGKF 1860
QY 1615 QVSAADDAQAKGKNNKSEDAKSVHESSTSESVLSSSPSTLVKEPNCITIPGISD 1674
Db 1861 QVSAADDAQAKGKNNKSEDAKSVHESSTSESVLSSSPSTLVKEPNCITIPGISD 1920
QY 1675 VPESAHKTTTAEASDTCQPTKVGAFQYTTTANKVGRSVKTEDEKIDTDTKEGPVAPSP 1734
Db 1921 VPESAHKTTTAEASDTCQPTKVGAFQYTTTANKVGRSVKTEDEKIDTDTKEGPVAPSP 1980
QY 1735 FMDLEQAVLPVAVTPKEKEPELSEPSHLGSPSSDPEAAFLSRVDGSGSPHSPOLSKS 1794
Db 1981 FMDLEQAVLPVAVTPKEKEPELSEPSHLGSPSSDPEAAFLSRVDGSGSPHSPOLSKS 2040
QY 1795 LPSQNLGSLNSFNSSYMSSDNEDJEDEDJELKELRLRDXKHKEIIDLQSRQKHIES 1854
Db 2041 LPSQNLGSLNSFNSSYMSSDNEDJEDEDJELKELRLRDXKHKEIIDLQSRQKHIES 2100
QY 1855 LYTLLGKVPVAVITPPAPPLSGRRRRPTKSGKSSRSSISGKNSPOLSGMLSGQSAASV 1914
Db 2101 LYTLLGKVPVAVITPPAPPLSGRRRRPTKSGKSSRSSISGKNSPOLSGMLSGQSAASV 2160
QY 1915 LHPQOTLHPGNIPESGONOLLOPKPSPSDNLVSATSPGATSVPSLSAPGOQTSSTN 1974
Db 2161 LHPQOTLHPGNIPESGONOLLOPKPSPSDNLVSATSPGATSVPSLSAPGOQTSSTN 2220
QY 1975 TVGATVNSQAAQAPPATTSRKGFTTDLKLIVDMNARMANLSGRRGSGKHNNYEPG 2034
Db 2221 TVGATVNSQAAQAPPATTSRKGFTTDLKLIVDMNARMANLSGRRGSGKHNNYEPG 2280
QY 2035 MARKFSAGQCLISMTSNLGSAPISASASATSLGHFTSMCPQOYGPATPFGQWNGGT 2094
Db 2281 MARKFSAGQCLISMTSNLGSAPISASASATSLGHFTSMCPQOYGPATPFGQWNGGT 2340
QY 2095 GGPAQPLGQFQPVGTASTLQNFNINSLQKISNPPGSLRTT 2136
Db 2341 GGPAQPLGQFQPVGTASTLQNFNINSLQKISNPPGSLRTT 2382

RESULT 4
US-09-854-856-30

; Sequence 30, Application US/09854856
; Patent No. 6541252
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade

; APPLICANT: Hilbun, Erin
; APPLICANT: Donoho, Gregory
; APPLICANT: Turner, C. Alexander Jr.
; TITLE OF INVENTION: No. 6541252el Human Kinases and Polynucleotides
; TITLE OF INVENTION: Encoding the Same
; FILE REFERENCE: LEX-0178-USA
; CURRENT APPLICATION NUMBER: US/09/854,856
; PRIOR FILING DATE: 2001-05-14
; PRIOR APPLICATION NUMBER: US 60/206,015
; NUMBER OF SEQ. ID NOS: 64
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ. ID NO 30
; LENGTH: 2108
; TYPE: PR
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(2108)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-09-854-856-30

Query Match 98.5%; Score 10646; DB 4; Length 2108;
Best Local Similarity 98.7%; Pred. No. 0;
Matches 2108; Conservative 0; Mismatches 0; Indels 28; Gaps 1;
QY 1 MSGGAAGKOSTPGSLFLSPAPAPKNGSSSDSVGKLGAAADAVTGRTEYRRRRHT 60
Db 1 MSGGAAGKOSTPGSLFLSPAPAPKNGSSSDSVGKLGAAADAVTGRTEYRRRRHT 60
QY 61 MDKDSRGAATTTTTHRRFRFRSVTCDSNAITALELPLGLPLSLPQPSIPAAVPOGAPPEH 120
Db 61 MDKDSRGAATTTTTHRRFRFRSVTCDSNAITALELPLGLPLSLPQPSIPAAVPOGAPPEH 120
QY 121 REETVITATSVQAQOPPAAPAEQAVAGABSTVSSSKDPVQPSLVGSKKEPP 180
Db 121 REETVITATSVQAQOPPAAPAEQAVAGABSTVSSSKDPVQPSLVGSKKEPP 180
QY 181 ARSGSGGSAKEPQEESSQOQDIELEETKAVGMSNDGRFLKPDIEIGRSGFKTVYGLD 240
Db 181 ARSGSGGSAKEPQEESSQOQDIELEETKAVGMSNDGRFLKPDIEIGRSGFKTVYGLD 240
QY 241 TETVEVAMCELODKLTKEERQRFKEEAEMLKGLQHPNIVRFYDSWSTYKGCIVLV 300
Db 241 TETVEVAMCELODKLTKEERQRFKEEAEMLKGLQHPNIVRFYDSWSTYKGCIVLV 300
QY 301 TELMTSGTLTKYLRKFRVMKIKVLRSCROILKGLQFLHTTPTIHRDLKCDNIFITGP 360
Db 301 TELMTSGTLTKYLRKFRVMKIKVLRSCROILKGLQFLHTTPTIHRDLKCDNIFITGP 360
QY 361 TGSVKIGDLGATLKRAAFKSVIGTPEFMAPEYEEKYDESUVVAFGMCMLEMATSEY 420
Db 361 TGSVKIGDLGATLKRAAFKSVIGTPEFMAPEYEEKYDESUVVAFGMCMLEMATSEY 420
QY 421 PYSECQNAAQIYRRVTSGVKPAFDKVAIPVKEIIEGCIIRONDERYSIDLLNHAFFQ 480
Db 421 PYSECQNAAQIYRRVTSGVKPAFDKVAIPVKEIIEGCIIRONDERYSIDLLNHAFFQ 480
QY 481 EETGVARELAEDGEKIALMLIRIDIKLKQKYDNEAIEFSPLDERVPEVDQEM 540
Db 481 EETGVARELAEDGEKIALMLIRIDIKLKQKYDNEAIEFSPLDERVPEVDQEM 540
QY 541 VESGVYEGDHKTAKAIKDRVSLIKRREOROLVREOEKKQEESSLKQOVQSSASQ 600
Db 541 VESGVYEGDHKTAKAIKDRVSLIKRREOROLVREOEKKQEESSLKQOVQSSASQ 600
QY 601 TGIKOLPSASTGIPASTTSASVSTQVEPEPEADQOQLOQOQPSISVLSGTVDSQG 660
Db 601 TGIKOLPSASTGIPASTTSASVSTQVEPEPEADQOQLOQOQPSISVLSGTVDSQG 660
QY 661 SSVFESRVSQOQTVSYSGSHEQASHSTGVGHIPSTVQAOQOPHGVYPPSSVQSGQG 720
Db 661 SSVFESRVSQOQTVSYSGSHEQASHSTGVGHIPSTVQAOQOPHGVYPPSSVQSGQG 720

QY 721 QPSSSILGVSSQPIQHPOOQOQIQTAPPOCTQVSLQSTSSSEATTAQPVQOPAP 780
 Db 714 -----QQGIQQTAPPOQIVQVSLQSTSSSEATTAQPVQOPAP 752
 QY 781 QVLPVNSAGKSTQGVSVAPAPVAVAPQATOPTTLASVDASNDVSGMSDGENV 840
 Db 753 QVLPVNSAGKSTQGVSVAPAPVAVAPQATOPTTLASVDASNDVSGMSDGENV 812
 QY 841 PSSGRHGRITTKHRYKRSVSRSHKETSRRPKRLILVSNKGRVVECOJETHNRKAVT 900
 Db 813 PSSGRHGRITTKHRYKRSVSRSHKETSRRPKRLILVSNKGRVVECOJETHNRKAVT 872
 QY 901 FKFDLDGNDPEEIIATIMVNDPILAIERESFVDQVREIIEKADENLSDVSEPEQDGL 960
 Db 873 FKFDLDGNDPEEIIATIMVNDPILAIERESFVDQVREIIEKADENLSDVSEPEQDGL 932
 QY 961 ESLQKODYGVSGSGOKLEGEFKOPIDASSMPQOIGITPSSLTQVVSAGRRFIVSPVES 1020
 Db 933 ESLQKODYGVSGSGOKLEGEFKOPIDASSMPQOIGITPSSLTQVVSAGRRFIVSPVES 992
 QY 1021 RLRESKVPSEIITVTVAASTAOSPGMNLSSHASSLSLQOAFSELRRACMTGEPNTAPNF 1080
 Db 993 RLRESKVPSEIITVTVAASTAOSPGMNLSSHASSLSLQOAFSELRRACMTGEPNTAPNF 1052
 QY 1081 SHTGPTPPVVPPLSSIAQVPTTAATAPVATSSPPMDISTSVYQSEVTVTEEGIAGV 1140
 Db 1053 SHTGPTPPVVPPLSSIAQVPTTAATAPVATSSPPMDISTSVYQSEVTVTEEGIAGV 1112
 QY 1141 ATSTGVVTSGLPIPEPVSESVLSSVSSITIPAVVISITSSPSLQVPTSTSEIIVSSTA 1200
 Db 1113 ATSTGVVTSGLPIPEPVSESVLSSVSSITIPAVVISITSSPSLQVPTSTSEIIVSSTA 1172
 QY 1201 LYPVTVSATSASAGSATAATGPKPAPVAVSQOAGSTVGTATLSVSTTSSFBSTAQLS 1260
 Db 1173 LYPVTVSATSASAGSATAATGPKPAPVAVSQOAGSTVGTATLSVSTTSSFBSTAQLS 1232
 QY 1261 IQSSSSTSTPLAETVAVVSAHSIDKTSHTSTTGLAFSLASAPSSSSSGAGVSYSIQPGG 1320
 Db 1233 IQSSSSTSTPLAETVAVVSAHSIDKTSHTSTTGLAFSLASAPSSSSSGAGVSYSIQPGG 1292
 QY 1321 LHPVTVSVASASTIFLEQOAGPTSTPLIPQVPSIPPLVQEPVAVPAVQOITLHNSQPAL 1380
 Db 1293 LHPVTVSVASASTIFLEQOAGPTSTPLIPQVPSIPPLVQEPVAVPAVQOITLHNSQPAL 1352
 QY 1381 LHPNPHHCEVDSDTQPKAPGIDIDIKLEKLSLSLSESSSSGAGASVLSLETSLVIES 1440
 Db 1353 LHPNPHHCEVDSDTQPKAPGIDIDIKLEKLSLSLSESSSSGAGASVLSLETSLVIES 1412
 QY 1441 TVTPGIPITTAVAPSKLITSTSTCLPPTNLPLGTVALPVTVPVPGQVSTPVSTTSGVK 1500
 Db 1413 TVTPGIPITTAVAPSKLITSTSTCLPPTNLPLGTVALPVTVPVPGQVSTPVSTTSGVK 1472
 QY 1501 PGTAPEKPELTAKPVLVGTGLPAGTLPSEQLPFPSPSILTQSOQPLEDDAQLRRTLSL 1560
 Db 1473 PGTAPEKPELTAKPVLVGTGLPAGTLPSEQLPFPSPSILTQSOQPLEDDAQLRRTLSL 1532
 QY 1561 EXITTVTSVAVPVSAAATTAITEAGTOPQKGVSVQKEPVLATSSGAVFMKGFQVSVAA 1620
 Db 1533 EXITTVTSVAVPVSAAATTAITEAGTOPQKGVSVQKEPVLATSSGAVFMKGFQVSVAA 1592
 QY 1621 DGAQKEGKNSEDAKSVHFESSSTSESSVLSSSPESSTLVKPEPAGITIPGISDVPESAH 1680
 Db 1593 DGAQKEGKNSEDAKSVHFESSSTSESSVLSSSPESSTLVKPEPAGITIPGISDVPESAH 1652
 QY 1681 KTTASEKASTQGPITKGRFOVTTANKVGRFVSYSKTEKITTDTKKEGPVAPSPFMLEQ 1740
 Db 1653 KTTASEKASTQGPITKGRFOVTTANKVGRFVSYSKTEKITTDTKKEGPVAPSPFMLEQ 1712
 QY 1741 AVLPVAVIPKKEKELSPSHLNGPSSDPEAAFLSRVDGSGS PHSHOULSSLSQNL 1800
 Db 1713 AVLPVAVIPKKEKELSPSHLNGPSSDPEAAFLSRVDGSGS PHSHOULSSLSQNL 1772

QY 1801 SOSLSNFSNSSYMSDNESEDIEDLKLRLRDKHLKEIODLOSROKHIEISLYTKLG 1860
 Db 1773 SOSLSNFSNSSYMSDNESEDIEDLKLRLRDKHLKEIODLOSROKHIEISLYTKLG 1832
 QY 1861 KVPVAVIIPPAAPLSGRRRRPTKSGSKSSRSSISLGNKSPOLSGNLSQGSAAVYVHPQOT 1920
 Db 1833 KVPVAVIIPPAAPLSGRRRRPTKSGSKSSRSSISLGNKSPOLSGNLSQGSAAVYVHPQOT 1892
 QY 1921 LHPGNIPESGONLLOPLKPSPESSDNLVSAFTSDGALSVESLSAPGGGTSSTNTVGAIV 1980
 Db 1893 LHPGNIPESGONLLOPLKPSPESSDNLVSAFTSDGALSVESLSAPGGGTSSTNTVGAIV 1952
 QY 1981 NSQAAQOPPAMTSRRKGTFTDILHKLVDNARARMNLGSRGSKGHNNYEGPGAKRFS 2040
 Db 1953 NSQAAQOPPAMTSRRKGTFTDILHKLVDNARARMNLGSRGSKGHNNYEGPGAKRFS 2012
 QY 2041 APGOLCISMTSNLGSAPISAAASATSLGHFTKSMCPPOQYFPATPFQAGVSGTGAPAPQ 2100
 Db 2013 APGOLCISMTSNLGSAPISAAASATSLGHFTKSMCPPOQYFPATPFQAGVSGTGAPAPQ 2072
 QY 2101 PLGQFOVGVGASLQNFNINLSNLOKSIINPPGSLNRTT 2136
 Db 2073 PLGQFOVGVGASLQNFNINLSNLOKSIINPPGSLNRTT 2108

RESULT 5
 US-09-854-856-24
 ; Sequence 24, Application US/09854856
 ; Patent No. 6541252
 ; GENERAL INFORMATION:
 ; APPLICANT: Walke, D. Wade
 ; APPLICANT: Hilbun, Erin
 ; APPLICANT: Donoho, Gregory
 ; APPLICANT: Turner, C. Alexander Jr.
 ; TITLE OF INVENTION: No. 6541252el Human Kinases and Polynucleotides
 ; FILE REFERENCE: LEX-0178-USA
 ; CURRENT APPLICATION NUMBER: US/09/854, 856
 ; PRIOR FILING DATE: 2001-05-14
 ; PRIOR APPLICATION NUMBER: US 60/206, 015
 ; NUMBER OF SEQ ID NOS: 64
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 24
 ; LENGTH: 2201
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: VARIANT
 ; LOCATION: (1)..(2201)
 ; OTHER INFORMATION: Xaa = Any Amino Acid
 US-09-854-856-24

Query Match 97.9%; Score 10589.5; DB 4; Length 2201;
 Best Local Similarity 94.6%; Pred. No. 0;
 Matches 2108; Conservative 0; Mismatches 0; Indels 121; Gaps 2;

QY 1 MSGAAAEKQSGSTGSLFLSPAPAPKNGSSSDSSVGEKLGAAADAVTGTTEYRRRRRT 60
 Db 1 MSGAAAEKQSGSTGSLFLSPAPAPKNGSSSDSSVGEKLGAAADAVTGTTEYRRRRRT 60
 QY 61 MDKDSGGAATTTTTHRRFRFRSVICDSNATALELPLSLPOPSIPAVVPQAPPEPH 120
 Db 61 MDKDSGGAATTTTTHRRFRFRSVICDSNATALELPLSLPOPSIPAVVPQAPPEPH 120
 QY 121 REETVATATNSQVAAQPPAAAPGEQAVGAPASTVPSSSTKDRPVQSOSVLSKEEPP 180
 Db 121 REETVATATNSQVAAQPPAAAPGEQAVGAPASTVPSSSTKDRPVQSOSVLSKEEPP 180
 QY 181 ARSGSGGSAKEQOERSQOQDDIELEFKAQVMSNDGRFLKFDIEIGRGSFTVYKGLD 240
 Db 181 ARSGSGGSAKEQOERSQOQDDIELEFKAQVMSNDGRFLKFDIEIGRGSFTVYKGLD 240

TITLE OF INVENTION: No. 6541252el Human Kinases and Polynucleotides
 TITLE OF INVENTION: Encoding the Same
 FILE REFERENCE: LEX-0178-USA
 CURRENT APPLICATION NUMBER: US/09/854,856
 CURRENT FILING DATE: 2001-05-14
 PRIOR APPLICATION NUMBER: US 60/206,015
 PRIOR FILING DATE: 2000-05-19
 NUMBER OF SEQ ID NOS: 64
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 18
 LENGTH: 2354
 TYPE: PRT
 ORGANISM: Homo sapiens
 FEATURE:
 NAME/KEY: VARIANT
 LOCATION: (1)...(2354)
 OTHER INFORMATION: Xaa = Any Amino Acid
 US-09-854-856-18

Query Match 97.2%; Score 10513; DB 4; Length 2354;
 Best Local Similarity 88.5%; Pred. No. 0;
 Matches 2108; Conservative 0; Mismatches 0; Indels 274; Gaps 2;

QY 1 MSGGAEEKOSTPGLSLFPAPAPKNGSSSDSSVGEKLGAAAADAVTGRTEYRRRRHT 60
 DB 1 MSGGAEEKOSTPGLSLFPAPAPKNGSSSDSSVGEKLGAAAADAVTGRTEYRRRRHT 60
 QY 61 MDKSRGAATTTTTHRRFFRRSVICDSNATALLPLGLPLPSPSIPAAVPOGAPPEPH 120
 DB 61 MDKSRGAATTTTTHRRFFRRSVICDSNATALLPLGLPLPSPSIPAAVPOGAPPEPH 120
 QY 121 REETVATATSOVAQOPPAAPGEQAVGAPSTVSTSKDRPVOPSLVSGKEEPP 180
 DB 121 REETVATATSOVAQOPPAAPGEQAVGAPSTVSTSKDRPVOPSLVSGKEEPP 180
 QY 121 REETVATATSOVAQOPPAAPGEQAVGAPSTVSTSKDRPVOPSLVSGKEEPP 180
 DB 121 REETVATATSOVAQOPPAAPGEQAVGAPSTVSTSKDRPVOPSLVSGKEEPP 180
 QY 181 ARSGSGGSAKEPOEERSQOQDDIELEETAYAVGMSNDGRFLKFDIEIGRGSFKTVKGLD 240
 DB 181 ARSGSGGSAKEPOEERSQOQDDIELEETAYAVGMSNDGRFLKFDIEIGRGSFKTVKGLD 240
 QY 241 TETTVAVAMELODRKLTGSRORFKEEAEMKGLQHPNIVRFDSMESTYKGGKCVLV 300
 DB 241 TETTVAVAMELODRKLTGSRORFKEEAEMKGLQHPNIVRFDSMESTYKGGKCVLV 300
 QY 301 TELMTSGTLTKYLRKFYMKIKVLRSMCROLKGLQELHTPTPIIRHDLKCDNIFITGP 360
 DB 301 TELMTSGTLTKYLRKFYMKIKVLRSMCROLKGLQELHTPTPIIRHDLKCDNIFITGP 360
 QY 301 TELMTSGTLTKYLRKFYMKIKVLRSMCROLKGLQELHTPTPIIRHDLKCDNIFITGP 360
 DB 301 TELMTSGTLTKYLRKFYMKIKVLRSMCROLKGLQELHTPTPIIRHDLKCDNIFITGP 360
 QY 361 TGSVKIGDLGLATLKRASFASVIGTPEFMAPEMEYEEKYDESVDVYAFGMCLEMATSEY 420
 DB 361 TGSVKIGDLGLATLKRASFASVIGTPEFMAPEMEYEEKYDESVDVYAFGMCLEMATSEY 420
 QY 421 PYSCQNAAOIYRRVTSGVKVPASFPDKVAIPVKEIIEGCIKRONDEKYSIKDLINAFQ 480
 DB 421 PYSCQNAAOIYRRVTSGVKVPASFPDKVAIPVKEIIEGCIKRONDEKYSIKDLINAFQ 480
 QY 481 EETGVARELAEDDEGEKIAIKMLRIDIKKAKYKONENAIERSFLERDVEDVAQEM 540
 DB 481 EETGVARELAEDDEGEKIAIKMLRIDIKKAKYKONENAIERSFLERDVEDVAQEM 540
 QY 541 VESGYVEGEGHKTMAKAIKDRVSLIKRRREQRLVREGEKKQEESSLKQOVESASQ 600
 DB 541 VESGYVEGEGHKTMAKAIKDRVSLIKRRREQRLVREGEKKQEESSLKQOVESASQ 600
 QY 601 TGTGOLPSASTGTPTASTASVSTOVEPEPEADQHQLOQOOPSTSVLSDGTVDGSG 660
 DB 601 TGTGOLPSASTGTPTASTASVSTOVEPEPEADQHQLOQOOPSTSVLSDGTVDGSG 660
 QY 661 SSVFTSERVSSQOQTVSGSQHEQAHSTGTVPBGHPISTVQAOSQPHGYVPSSVAQSGSG 720
 DB 661 SSVFTSERVSSQOQTVSGSQHEQAHSTGTVPBGHPISTVQAOSQPHGYVPSSVAQSGSG 720
 QY 721 QPSSSSILTVSSSQPIQHPQOQOQIGTAPDQOQTVQVSLSTSTSSSEATTAQPVQOPAP 780
 DB 721 QPSSSSILTVSSSQPIQHPQOQOQIGTAPDQOQTVQVSLSTSTSSSEATTAQPVQOPAP 780

DB 714 -----QOQIQOQAPQOQVQVSLSTSTSSSEATTAQPVQOPAP 752
 QY 781 QVLPQVSAKQ----- 791
 DB 753 QVLPQVSAKQLPVSQPVPTIQGEPQIPVATQPSVVPVSHAHFLPVQQLPPTLLPOYP 812
 QY 792 ----- 791
 DB 813 VSQIPITPHVSTAQGTGSSLPITMAAGITQPLTLASSATTAAIPGVSTVPSQLPTLL 872
 QY 792 ----- 791
 DB 873 QPVTQLPSPQVHPQLQPAVQSMGIPANLQGAEBVPLSSGDVLVQGFPPRLPPOYPGDSNI 932
 QY 792 ----- 791
 DB 933 APSSNVAVCIHSTVLXPMPTPEVLATPPGYEPTVQPYVESNLLVPMGVGVQVQVSPG 992
 QY 792 -----STOGVQVAPAEVAVVQOPATOPTTLASSVDSASHSVASGMS 834
 DB 993 GSIAQATTTSSQOAVLESTQGVQVAPAEVAVVQOPATOPTTLASSVDSASHSVASGMS 1052
 QY 835 DGNENVSSSGRHEGRITTKHRYKSVASRSRHEKTSRPKRLINLVSNKGRVVECOLETH 894
 DB 1053 DGNENVSSSGRHEGRITTKHRYKSVASRSRHEKTSRPKRLINLVSNKGRVVECOLETH 1112
 QY 895 NRKMTTFKPLDDGNPEEITATIMNDFILAEBSVYDQREILEKADMLSDVSEY 954
 DB 1113 NRKMTTFKPLDDGNPEEITATIMNDFILAEBSVYDQREILEKADMLSDVSEY 1172
 QY 955 EGDQGLESLQKNDYGVSGSKLEGEFKOPIPASMSQOQIGIPSSLTQVHSGRPFIV 1014
 DB 1173 EGDQGLESLQKNDYGVSGSKLEGEFKOPIPASMSQOQIGIPSSLTQVHSGRPFIV 1222
 QY 1015 SPVSESLRESKVPSEIITDVAASTAQSPQMLSHSASSLSQOAFSELRRAQMTGPN 1074
 DB 1233 SPVSESLRESKVPSEIITDVAASTAQSPQMLSHSASSLSQOAFSELRRAQMTGPN 1292
 QY 1075 TAPNPSHTGPTPVVPEPLSSTAGVPTTAAATPVPATSSPPNDISTSVYQSVYVTE 1134
 DB 1293 TAPNPSHTGPTPVVPEPLSSTAGVPTTAAATPVPATSSPPNDISTSVYQSVYVTE 1352
 QY 1135 EGIAGVATSTGVTSGLPIPPSESPVLSVSSSIIPVAVSISTSPSIQVPTSEI 1194
 DB 1353 EGIAGVATSTGVTSGLPIPPSESPVLSVSSSIIPVAVSISTSPSIQVPTSEI 1412
 QY 1195 VVSSSTALVPSVTVSATSASAGSSTATPGPKPAPVAVSQAAAGSTTVGATLTVSVSTTSPS 1254
 DB 1413 VVSSSTALVPSVTVSATSASAGSSTATPGPKPAPVAVSQAAAGSTTVGATLTVSVSTTSPS 1472
 QY 1255 TAOQLSTQLSSSTPTLATTVVVAHSLDKTSHSSTTGIAFLSAPSSSSPQAGVSSY 1314
 DB 1473 TAOQLSTQLSSSTPTLATTVVVAHSLDKTSHSSTTGIAFLSAPSSSSPQAGVSSY 1532
 QY 1315 ISOQGLHPVLPVSVIASTPILPQAAGPTSTPLLPQVPSIPVQVPAANTPAVQOQLIHS 1374
 DB 1533 ISOQGLHPVLPVSVIASTPILPQAAGPTSTPLLPQVPSIPVQVPAANTPAVQOQLIHS 1592
 QY 1375 QOPPALIPNOPTHCHPEVSDTOPKAPGIDDIKTEBKLSLFSHSSSQAQHAASVLET 1434
 DB 1593 QOPPALIPNOPTHCHPEVSDTOPKAPGIDDIKTEBKLSLFSHSSSQAQHAASVLET 1652
 QY 1435 SLVIESVTPGIPPTVAVPSKLTSTSTSLPPTNPLPGVVALPVMVYVTPGVSTPVST 1494
 DB 1653 SLVIESVTPGIPPTVAVPSKLTSTSTSLPPTNPLPGVVALPVMVYVTPGVSTPVST 1712
 QY 1495 TTSQVKGTPASKRPPLTKAVLPLVGTLPAGTLPSEQLPFPGPSLTQSQOPELDDACL 1554
 DB 1713 TTSQVKGTPASKRPPLTKAVLPLVGTLPAGTLPSEQLPFPGPSLTQSQOPELDDACL 1772
 QY 1555 RRTLSPBXITTVTSAGVSVMAAPTAITEAGTOPQKGVQVKEGVPVLTSSGAGVFXKGRF 1614
 DB 1773 RRTLSPBXITTVTSAGVSVMAAPTAITEAGTOPQKGVQVKEGVPVLTSSGAGVFXKGRF 1832

QY 1615 QVSVAADGAQKGNKSEDAKSVHFEESSTESSVLSSSSPESTLYKEPENGITIPGISSD 1674
DB 1833 QVSVAADGAQKGNKSEDAKSVHFEESSTESSVLSSSSPESTLYKEPENGITIPGISSD 1892
QY 1675 VPBSAHKTTSAEASDLTGQPTKVRFOYTTANKVGFVSVKTEBDKITDTIKKEGPVAPSP 1734
DB 1693 VPBSAHKTTSAEASDLTGQPTKVRFOYTTANKVGFVSVKTEBDKITDTIKKEGPVAPSP 1952
QY 1735 FMDLEQAVLPVAVIIPKKEKPELSEPSHNGPESDPEAFLSRDVEDGSGSPSPHQLSKS 1794
DB 1953 FMDLEQAVLPVAVIIPKKEKPELSEPSHNGPESDPEAFLSRDVEDGSGSPSPHQLSKS 2012
QY 1795 LPSQNLQSLSNSFNSSNSDNESEDIEDLKLRLRDKHKEIODLQSRKHIES 1854
DB 2013 LPSQNLQSLSNSFNSSNSDNESEDIEDLKLRLRDKHKEIODLQSRKHIES 2072
QY 1855 LYTGLGVPPAVIIPPAAPLSGRRRRPTKSGKSSSSSGKNSPOLSGNLSGQSAASV 1914
DB 2073 LYTGLGVPPAVIIPPAAPLSGRRRRPTKSGKSSSSSGKNSPOLSGNLSGQSAASV 2132
QY 1915 LHPQOQLHPGNIPESSQNLQPLKPSPSDNLVSAFTSDGATSVPSLSAPGQGTSTN 1974
DB 2133 LHPQOQLHPGNIPESSQNLQPLKPSPSDNLVSAFTSDGATSVPSLSAPGQGTSTN 2192
QY 1975 TVGATVNSQAAQAPRPMATSSRKCTFTDHLKLVDMNARDAMNLSGRGSGKHNNYEGPG 2034
DB 2193 TVGATVNSQAAQAPRPMATSSRKCTFTDHLKLVDMNARDAMNLSGRGSGKHNNYEGPG 2252
QY 2035 MARKFSAAPGOLCISMTSNLGGAPISASAATSLGHFTKSCWCPQOYGFPAIPFGAOMSGT 2094
DB 2253 MARKFSAAPGOLCISMTSNLGGAPISASAATSLGHFTKSCWCPQOYGFPAIPFGAOMSGT 2312
QY 2095 GGPAPOPLGQFOPVGTASLQNFENISNLOKISNPPGSLTKTT 2136
DB 2313 GGPAPOPLGQFOPVGTASLQNFENISNLOKISNPPGSLTKTT 2354

RESULT 7
US-09-854-856-46
Sequence 46, Application US/09854856
Patent No. 6541252
GENERAL INFORMATION:
APPLICANT: Walke, D. Wade
APPLICANT: Hilbun, Erin
APPLICANT: Donoho, Gregory
APPLICANT: Turner, C. Alexander Jr.
TITLE OF INVENTION: No. 6541252el Human Kinases and Polynucleotides
FILE REFERENCE: Lex-0178-USA
CURRENT FILING DATE: 2001-05-14
PRIORITY FILING DATE: 2000-05-19
NUMBER OF SEQ ID NOS: 64
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 46
LENGTH: 2076
TYPE: PRN
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: VARIANT
LOCATION: (1)...(2076)
OTHER INFORMATION: Xaa = Any Amino Acid
US-09-854-856-46

Query Match 97.2%; Score 10508; DB 4; Length 2076;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2076; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 61 MDKSRGAATTTTTHFRFRSVICDGNATALELPGPLSLPQPSIPAAVQASAPPEPH 120
DB 1 MDKSRGAATTTTTHFRFRSVICDGNATALELPGPLSLPQPSIPAAVQASAPPEPH 60

QY 121 REETVATATATQVAQOQPPAAABEGQAVAGPAPSTVPSSTSKDRPVSQPSLVGSKPEPP 180
DB 61 REETVATATATQVAQOQPPAAABEGQAVAGPAPSTVPSSTSKDRPVSQPSLVGSKPEPP 120
QY 181 ARSGGGGSAKEPQEBRSQOQDDIIELETAVKMSNDGRFLKDEIEGRGSPKTVYGLD 240
DB 121 ARSGGGGSAKEPQEBRSQOQDDIIELETAVKMSNDGRFLKDEIEGRGSPKTVYGLD 180
QY 241 TETTVAVAMELODRKLTTSEROPFKREAMELGLQHPNIVRFDSVESTYKGGKCVLV 300
DB 181 TETTVAVAMELODRKLTTSEROPFKREAMELGLQHPNIVRFDSVESTYKGGKCVLV 240
QY 301 TELMTSGTLTKYKRFKVMKI KVLRSWCROILKGLQPLHRTPTPIIHRDLKCDNIFITGP 360
DB 241 TELMTSGTLTKYKRFKVMKI KVLRSWCROILKGLQPLHRTPTPIIHRDLKCDNIFITGP 300
QY 361 TGSVKIGDLGLATLKRASPAKSVYGTPEFMAPEMEYBEKYBSVDVYAFGCMLEMATSEY 420
DB 301 TGSVKIGDLGLATLKRASPAKSVYGTPEFMAPEMEYBEKYBSVDVYAFGCMLEMATSEY 360
QY 421 PYSECQNAQIYRRVTSGVBPASFDKVAIPEVKEIIEGCI RQNDREYSIKDLINHAFFQ 480
DB 361 PYSECQNAQIYRRVTSGVBPASFDKVAIPEVKEIIEGCI RQNDREYSIKDLINHAFFQ 420
QY 481 EETGVAVELAEEDGEXIAIKLWLRIEDIKLKGXYKDNDAIEBSFDLBERDVEDVAQEM 540
DB 421 EETGVAVELAEEDGEXIAIKLWLRIEDIKLKGXYKDNDAIEBSFDLBERDVEDVAQEM 480
QY 541 VESGYVCEGHNKTMAKIKORVSLIKRKROROLVREOEKKYKOEBSLKQOVBQSSAQ 600
DB 481 VESGYVCEGHNKTMAKIKORVSLIKRKROROLVREOEKKYKOEBSLKQOVBQSSAQ 540
QY 601 TGIKOLPSASTGIPATSTASVSTQVEPEPEADHQOLQYQOPSTSVLSDGVNDSGQ 660
DB 541 TGIKOLPSASTGIPATSTASVSTQVEPEPEADHQOLQYQOPSTSVLSDGVNDSGQ 600
QY 661 SSVFTEBSRVSSQOTVSYGSOHEQAHSTGTVPGHIPSTVQAQSQPHGVYPPSSVVAQSQSQ 720
DB 601 SSVFTEBSRVSSQOTVSYGSOHEQAHSTGTVPGHIPSTVQAQSQPHGVYPPSSVVAQSQSQ 660
QY 721 QPSSSLTGVSSSQPIHPQOQOQIGQTPAPQOTVQVSLQSTSTSSAATTAQVSPQAP 780
DB 661 QPSSSLTGVSSSQPIHPQOQOQIGQTPAPQOTVQVSLQSTSTSSAATTAQVSPQAP 720
QY 781 QVLPQVAGKQSTQGVQVAPAEVAVAOQATQPTTLASVDSABSDVAGSDGMBNV 840
DB 721 QVLPQVAGKQSTQGVQVAPAEVAVAOQATQPTTLASVDSABSDVAGSDGMBNV 780
QY 841 PSSSGRHEGRTTKRHKYSVRSRHEKTSRPKLILVNSNKGDRVVECOLETHNRKVT 900
DB 781 PSSSGRHEGRTTKRHKYSVRSRHEKTSRPKLILVNSNKGDRVVECOLETHNRKVT 840
QY 901 FKFDLDQNPBEIATIVNNDPILATRESFVQVRIIEKADMLESDVSVBEGQGL 960
DB 841 FKFDLDQNPBEIATIVNNDPILATRESFVQVRIIEKADMLESDVSVBEGQGL 900
QY 961 ESLOGKDDYGPSSQKLEGEFKQPIPASSMPQOIGITSSLTQVYHSGRRFVSPBES 1020
DB 901 ESLOGKDDYGPSSQKLEGEFKQPIPASSMPQOIGITSSLTQVYHSGRRFVSPBES 960
QY 1021 RLRESKVPSEBITDTVAASTAQSFGNMLSHSASSLSQQAFFSELRRQMTGPTAPPNF 1080
DB 961 RLRESKVPSEBITDTVAASTAQSFGNMLSHSASSLSQQAFFSELRRQMTGPTAPPNF 1020
QY 1081 SHTGPTFPVVPVPPFSSITAGVPTTAATAFPATSSPNDISTSVIOSEVYPTBEGTAGV 1140
DB 1021 SHTGPTFPVVPVPPFSSITAGVPTTAATAFPATSSPNDISTSVIOSEVYPTBEGTAGV 1080
QY 1141 ATSTGVVTSGLPLPPIVSESPVLSVYSITIPVVISITSPSLQVPTSBIVSSTA 1200
DB 1081 ATSTGVVTSGLPLPPIVSESPVLSVYSITIPVVISITSPSLQVPTSBIVSSTA 1140

```

QY 1201 LPSVTSASASAGSTATGPKPRAVVSQAAGSTTVGATLTSTSTTSFSPSTASQLS 1260
Db 1141 LPSVTSASASAGSTATGPKPRAVVSQAAGSTTVGATLTSTSTTSFSPSTASQLS 1200
QY 1261 IOLSSSTSTPLTAETVVSANSLDKTSHSTTGALFSLASAPSSSSPGAGVSSYISQPGG 1330
Db 1201 IOLSSSTSTPLTAETVVSANSLDKTSHSTTGALFSLASAPSSSSPGAGVSSYISQPGG 1260
QY 1321 LHPVIVPVIASSTPLIPQAGPTSTPLIPQVPSIPVIVQVAVNPVAVQOTLIHSQOPAL 1380
Db 1261 LHPVIVPVIASSTPLIPQAGPTSTPLIPQVPSIPVIVQVAVNPVAVQOTLIHSQOPAL 1320
QY 1381 LHPVIVPVIASSTPLIPQAGPTSTPLIPQVPSIPVIVQVAVNPVAVQOTLIHSQOPAL 1440
Db 1321 LHPVIVPVIASSTPLIPQAGPTSTPLIPQVPSIPVIVQVAVNPVAVQOTLIHSQOPAL 1380
QY 1441 TVTGTGIPPTAVAPSKLTSTSTSTCLPPTNLPLGVVALPVTVPVTPGVSTVSTTSQV 1500
Db 1381 TVTGTGIPPTAVAPSKLTSTSTSTCLPPTNLPLGVVALPVTVPVTPGVSTVSTTSQV 1440
QY 1501 PGTAAPKPLTKAVLPVGTLPAGTLPSBQLPFPQPSLTQSOQPLEDLAOLRLTSP 1560
Db 1441 PGTAAPKPLTKAVLPVGTLPAGTLPSBQLPFPQPSLTQSOQPLEDLAOLRLTSP 1500
QY 1561 EXITVTSAVGVMAAPTAITEAGTOPQGVSYQVKEGVLATSSGAGVFKMGAFQVSYAA 1620
Db 1501 EXITVTSAVGVMAAPTAITEAGTOPQGVSYQVKEGVLATSSGAGVFKMGAFQVSYAA 1560
QY 1621 DGAKGKGNKSEDAKSVHFEESTSESSVLSSESTLVKPEPNCITIPGISVVPESAH 1680
Db 1561 DGAKGKGNKSEDAKSVHFEESTSESSVLSSESTLVKPEPNCITIPGISVVPESAH 1620
QY 1681 KTTSSEAKSDTGTQTKGRFOVTTTANKVGRFYSKTEDKITDPRKKGPAVSPFMLEQ 1740
Db 1621 KTTSSEAKSDTGTQTKGRFOVTTTANKVGRFYSKTEDKITDPRKKGPAVSPFMLEQ 1680
QY 1741 AVLEPAVIPKKEPELSPSHLNGPSSDPEAFSLRDVDDGSGSPHSPHOLSKSLPSQNL 1800
Db 1681 AVLEPAVIPKKEPELSPSHLNGPSSDPEAFSLRDVDDGSGSPHSPHOLSKSLPSQNL 1740
QY 1801 SOSLSNFPNSYMSDNESDIEDDLKLELRDLADKHLKEI QDI OSQKHEIESLYTKLG 1860
Db 1741 SOSLSNFPNSYMSDNESDIEDDLKLELRDLADKHLKEI QDI OSQKHEIESLYTKLG 1800
QY 1861 KVPRAVILIPPAAPISGRRRRTKSGKSSRSLSGNKSPOLSGNLSGQSAASVLAHQOT 1920
Db 1801 KVPRAVILIPPAAPISGRRRRTKSGKSSRSLSGNKSPOLSGNLSGQSAASVLAHQOT 1860
QY 1921 LHPGNIPESSGQNLQPLKPSPSDNLYSAFTSDGAI SVPSLSAPQOGTSTTVGATV 1980
Db 1861 LHPGNIPESSGQNLQPLKPSPSDNLYSAFTSDGAI SVPSLSAPQOGTSTTVGATV 1920
QY 1981 NSQAQAQOPAMTSSRRGTFTDHLKLVDMWAPAMNLSGRGSKGMNTEGPMARKFS 2040
Db 1921 NSQAQAQOPAMTSSRRGTFTDHLKLVDMWAPAMNLSGRGSKGMNTEGPMARKFS 1980
QY 2041 APQOLCISMTSNLGSAPI SAASATSLGHTKSCPCPOQYGFPAIPFGAOMSGGAPQ 2100
Db 1981 APQOLCISMTSNLGSAPI SAASATSLGHTKSCPCPOQYGFPAIPFGAOMSGGAPQ 2040
QY 2101 PLGQFQVGTASLQNFNISNLQKISINPFGSNLRTT 2136
Db 2041 PLGQFQVGTASLQNFNISNLQKISINPFGSNLRTT 2076

```

```

; APPLICANT: Turner, C. Alexander Jr.
; TITLE OF INVENTION: No. 6541252el Human Kinases and Polynucleotides
; FILE REFERENCE: LEX-0178-USA
; CURRENT FILING DATE: 2001-05-14
; PRIOR FILING DATE: 2000-05-19
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 40
; LENGTH: 2169
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: VARIANT
; LOCATION: (1) (2169)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-09-854-856-40

Query Match 96.7%; Score 10451.5; DB 4; Length 2169;
Beet Local Similarity 95.7%; Pred. No. 0;
Matches 2076; Conservative 0; Mismatches 0; Indels 93; Gaps 1;

QY 61 MDKDSRGAATTTTTEHFRFRSVYICDSNATALELPGILPQPSIPAAVPOSAPPEPH 120
Db 1 MDKDSRGAATTTTTEHFRFRSVYICDSNATALELPGILPQPSIPAAVPOSAPPEPH 60

QY 121 REETVTAATSOVAQOPPAAPGEQAVAPAPSTVBSSTKDRPVASQPSIVSGKEPPP 180
Db 61 REETVTAATSOVAQOPPAAPGEQAVAPAPSTVBSSTKDRPVASQPSIVSGKEPPP 120

QY 181 ARSGSGGSAKEQPEESQOQDDIEELTAQVGSNDGRFLKPIETIGRSFKTVYGLD 240
Db 121 ARSGSGGSAKEQPEESQOQDDIEELTAQVGSNDGRFLKPIETIGRSFKTVYGLD 180

QY 241 TETTVAVAKELQDRKTKSEORFKEEAMLKGLQHPNIVRFYDSNESTVKGKCIYLV 300
Db 181 TETTVAVAKELQDRKTKSEORFKEEAMLKGLQHPNIVRFYDSNESTVKGKCIYLV 240

QY 301 TELMTSGTLKTYLKRKVMKIKVLRNCRQILKGLQPLHTRTPPIIHRDLKCDNI FTGP 360
Db 241 TELMTSGTLKTYLKRKVMKIKVLRNCRQILKGLQPLHTRTPPIIHRDLKCDNI FTGP 300

QY 361 TGSVKIGDGLATLTKASFAKSVITGPEFMAPEVBEKXDESVDVYAFGCMEMATSEY 420
Db 301 TGSVKIGDGLATLTKASFAKSVITGPEFMAPEVBEKXDESVDVYAFGCMEMATSEY 360

QY 421 PYSECQAAQIYRRVTSVGPAPAFDVAIPEVKEIIEGCI RQNDERYSIKDLNHAFFQ 480
Db 361 PYSECQAAQIYRRVTSVGPAPAFDVAIPEVKEIIEGCI RQNDERYSIKDLNHAFFQ 420

QY 481 EETGVVEALEEDDEKIAIKMLRIEDIKLKGXYKXNEAIEFSPDLERVPEDVAQEM 540
Db 421 EETGVVEALEEDDEKIAIKMLRIEDIKLKGXYKXNEAIEFSPDLERVPEDVAQEM 480

QY 541 VESGVYCEGHKTMARAIKQVSLIKRKEQRLVBEQKQOEBSLKOQVEQSSASQ 600
Db 481 VESGVYCEGHKTMARAIKQVSLIKRKEQRLVBEQKQOEBSLKOQVEQSSASQ 540

QY 601 TGIKOLPASASTGIPTASTTSASVSTQVEPEPADHQOLQYOOPSISVLSDTGVSQGG 660
Db 541 TGIKOLPASASTGIPTASTTSASVSTQVEPEPADHQOLQYOOPSISVLSDTGVSQGG 600

QY 661 SSVFTESRVSSQQTIVSGQHEQAHSTGVPHI PSTVQAQOSOPHGVYPPSSVAQOSQG 720
Db 601 SSVFTESRVSSQQTIVSGQHEQAHSTGVPHI PSTVQAQOSOPHGVYPPSSVAQOSQG 660

QY 721 QPSSSSLTGVSSSOPLOHPQOQOQIOQTAPPOOTVOYYSQTSSEATTAQVSOQAP 780
Db 661 QPSSSSLTGVSSSOPLOHPQOQOQIOQTAPPOOTVOYYSQTSSEATTAQVSOQAP 720

QY 781 QVLPVVSAGKO----- 791

```

RESULT 8
 US-09-854-856-40
 ; Sequence 40; Application US/09854856
 ; Patent No. 6541252
 ; GENERAL INFORMATION:
 ; APPLICANT: Walke, D. Wade
 ; APPLICANT: Hilbun, Erin
 ; APPLICANT: Donoho, Gregory

```
Db 721 QVLPOVSAGKGFPPRLPQPGDSNIAPISSNVASVCIHSTVLXPMPTLEVLATPGYFPT 780
Qy 792 -----STQGVQVAPAEVAV 867
Db 781 VVQPVYESNLLVPMGVGVQGVQVVSQPGGSLAQAPITSSQOAVLESTQGVQVAPAEVAV 840
Qy 808 ACPQATQPTTLASSVDSAHSDVASGMSDGNENVPSSSGRHEGRTTKRHKYKVSRSRHE 867
Db 841 ACPQATQPTTLASSVDSAHSDVASGMSDGNENVPSSSGRHEGRTTKRHKYKVSRSRHE 900
Qy 868 KTSRPKLRILNVSNKGDVVECOLETHNRMTVTFKFDLDGDNBEIATIMWNNDFILAI 927
Db 901 KTSRPKLRILNVSNKGDVVECOLETHNRMTVTFKFDLDGDNBEIATIMWNNDFILAI 960
Qy 928 RESFVQVREIIEKADENLSEDSVVEPEGQGLSELOGKXDYGPSQKLEGEFKQPIPA 987
Db 961 RESFVQVREIIEKADENLSEDSVVEPEGQGLSELOGKXDYGPSQKLEGEFKQPIPA 1020
Qy 988 SSMPOQIGITSSLTQVHSGRRFIVSPVESRLRESKYFSEITDTVAASTAQSPGM 1047
Db 1021 SSMPOQIGITSSLTQVHSGRRFIVSPVESRLRESKYFSEITDTVAASTAQSPGM 1080
Qy 1048 LSHSASSLSLQOAFSEILRAQMTGENTADPNFSHTGPTPVVPPFLSLAGVPTTAAT 1107
Db 1081 LSHSASSLSLQOAFSEILRAQMTGENTADPNFSHTGPTPVVPPFLSLAGVPTTAAT 1140
Qy 1108 APVATSSPNDISTSVIOSEVTVTEEGTAGVATSGVTSQGLPIPVSESPVLSVV 1167
Db 1141 APVATSSPNDISTSVIOSEVTVTEEGTAGVATSGVTSQGLPIPVSESPVLSVV 1200
Qy 1168 SSITIPAVSISTSPLOVPTSEIIVSSTALPVSATVATSAAGSSTATPGPKPA 1227
Db 1201 SSITIPAVSISTSPLOVPTSEIIVSSTALPVSATVATSAAGSSTATPGPKPA 1260
Qy 1228 VVSQQAAGSTTVGATLTSVSTTSPSTASQSLTQLSSSTPTPLAFTVVVAHSLDKTS 1287
Db 1261 VVSQQAAGSTTVGATLTSVSTTSPSTASQSLTQLSSSTPTPLAFTVVVAHSLDKTS 1320
Qy 1288 HSSITGLAFSLASASSSSPGAGSVSIQPGGHPVIVSVASTIILQOAGPTSTPL 1347
Db 1321 HSSITGLAFSLASASSSSPGAGSVSIQPGGHPVIVSVASTIILQOAGPTSTPL 1380
Qy 1348 LPQVPSIPPLVQVAVNPVAVOQTLIHQOPALLPNOPTHCPVEDSDTPKAPGIDIK 1407
Db 1381 LPQVPSIPPLVQVAVNPVAVOQTLIHQOPALLPNOPTHCPVEDSDTPKAPGIDIK 1440
Qy 1408 TLEBKRLSLFSEHSSGAGAHASVLETSVLESTVTPGIPPTAVAPSKLLTSTTSTCLP 1467
Db 1441 TLEBKRLSLFSEHSSGAGAHASVLETSVLESTVTPGIPPTAVAPSKLLTSTTSTCLP 1500
Qy 1468 TNLPLGLVALPVTVPVTPGVSTVSTTSGVKRGTAHSKRPRLTKAVLVPVGTLEPLAGTL 1527
Db 1501 TNLPLGLVALPVTVPVTPGVSTVSTTSGVKRGTAHSKRPRLTKAVLVPVGTLEPLAGTL 1560
Qy 1528 PSEOLPFPFPGSLTQSOQPLEDDAOLRRTLSPEXITVTSAGVGVMAAPTAITEAGTOP 1587
Db 1561 PSEOLPFPFPGSLTQSOQPLEDDAOLRRTLSPEXITVTSAGVGVMAAPTAITEAGTOP 1620
Qy 1588 QKGVSYQKEGVLATSSGAGVFKMGRFQVSVAAQAKGEGKNKSEDAKSVHESSTSESS 1647
Db 1621 QKGVSYQKEGVLATSSGAGVFKMGRFQVSVAAQAKGEGKNKSEDAKSVHESSTSESS 1680
Qy 1648 VLSASSSESTLVKEPEPGITRGISDVPEBAHKTSEAKSDPGQPKVCRPOVTTAN 1707
Db 1681 VLSASSSESTLVKEPEPGITRGISDVPEBAHKTSEAKSDPGQPKVCRPOVTTAN 1740
Qy 1708 KVGFFSVSKTEDKLTDTKKKGVPVAPPFMDLEQAVLPAVLPKKEKPELSEPSHANGSSD 1767
Db 1741 KVGFFSVSKTEDKLTDTKKKGVPVAPPFMDLEQAVLPAVLPKKEKPELSEPSHANGSSD 1800
Qy 1768 PEAAPLSBDVDSGSGSPHSPHOLSSKSLPSQNLSSNSFSSNGSSDNESEDIEDDLK 1827
```

```
Db 1801 PEAAPLSBDVDSGSGSPHSPHOLSSKSLPSQNLSSNSFSSNGSSDNESEDIEDDLK 1860
Qy 1828 LELRLRDLKHEIIOLOSRQKHEIESTYTKLGKVPVAVIIPPAADLSGRRRPTSKGS 1887
Db 1861 LELRLRDLKHEIIOLOSRQKHEIESTYTKLGKVPVAVIIPPAADLSGRRRPTSKGS 1920
Qy 1888 KSSRSSSLGNKSPQLSGNISGOSASVLHPQOTLHPGNIPESGQNLLOPLKPSPSDN 1947
Db 1921 KSSRSSSLGNKSPQLSGNISGOSASVLHPQOTLHPGNIPESGQNLLOPLKPSPSDN 1980
Qy 1948 LYSAPTSDAIIVPSLSAPCGQSTSTNTYGATNLSQAQAQAPAMTSSRKGTEDDLHKL 2007
Db 1981 LYSAPTSDAIIVPSLSAPCGQSTSTNTYGATNLSQAQAQAPAMTSSRKGTEDDLHKL 2040
Qy 2008 VDMWADAMNLSGRSGSKGHMNYEGGMARKFAPOLCISMTSNLGSAPISASATSL 2067
Db 2041 VDMWADAMNLSGRSGSKGHMNYEGGMARKFAPOLCISMTSNLGSAPISASATSL 2100
Qy 2068 GHFTKSCMPQOYGFPAFPFGAQWSGTGAPQPLQOPQFVGTASIQNENISMLQKISN 2127
Db 2101 GHFTKSCMPQOYGFPAFPFGAQWSGTGAPQPLQOPQFVGTASIQNENISMLQKISN 2160
Qy 2128 PGGSNLRIT 2136
Db 2161 PGGSNLRIT 2169

RESULT 9
US-09-854-856-34
; Sequence 34, Application US/09854856
; Patent No. 6541252
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Hillman, Erin
; APPLICANT: Donoho, Gregory
; APPLICANT: Turner, C. Alexander Jr.
; TITLE OF INVENTION: No. 6541252el Human Kinases and polynucleotides
; TITLE OF INVENTION: Encoding the Same
; FILE REFERENCE: Lex-0178-USA
; CURRENT APPLICATION NUMBER: US/09/854,856
; PRIOR FILING DATE: 2001-05-14
; PRIOR APPLICATION NUMBER: US 60/206,015
; PRIOR FILING DATE: 2000-05-19
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 34
; LENGTH: 2322
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)-(2322)
; OTHER INFORMATION: xaa = Any Amino Acid
US-09-854-856-34

Query Match 96.0%; Score 10375; DB 4; Length 2322;
Best Local Similarity 89.4%; Pred. No. 0;
Matches 2076; Conservative 0; Mismatches 0; Indels 246; Gaps 1;
```

Db 181 TETTEVAMCELODRKLTSEORKEBEAMLKGLQHPNIVRFYDWSWETVKGKCCIVLV 240
 QY 301 TELMTSGTLKTVLRFKFMKTKVLSMGRQLKGIQFHTTTPPIIHBDLKCDNIFITGP 360
 Db 241 TELMTSGTLKTVLRFKFMKTKVLSMGRQLKGIQFHTTTPPIIHBDLKCDNIFITGP 300
 QY 361 TGSVATIGDLATLKRASFASVIGTPEPFMAPEMYEEXYDESVDVYAGMCMLEATSEY 420
 Db 301 TGSVATIGDLATLKRASFASVIGTPEPFMAPEMYEEXYDESVDVYAGMCMLEATSEY 360
 QY 421 PYSECQNAAOIYRYRYTSVGRKASPDKVAIPVKEIIEGCIPIKONDERYSIKDNLNHAFO 480
 Db 361 PYSECQNAAOIYRYRYTSVGRKASPDKVAIPVKEIIEGCIPIKONDERYSIKDNLNHAFO 420
 QY 481 EETGVRELAEBEDGEEKIAIKLMLRIEDIKLKKGKYNONEAIEFSFDLERVPEPDVAQEM 540
 Db 421 EETGVRELAEBEDGEEKIAIKLMLRIEDIKLKKGKYNONEAIEFSFDLERVPEPDVAQEM 480
 QY 541 VESGYVEGDHKTMAKIKORVSLIKRREOROLVREOEKKOBEESLKQOEVASASQ 600
 Db 481 VESGYVEGDHKTMAKIKORVSLIKRREOROLVREOEKKOBEESLKQOEVASASQ 540
 QY 601 TGIKOLASASTGIFTASTSASTOYEPPEPEADHQOLOYOOPSISVLSGTVDSGQG 660
 Db 541 TGIKOLASASTGIFTASTSASTOYEPPEPEADHQOLOYOOPSISVLSGTVDSGQG 600
 QY 661 SSVFTESRVSQOQTVSYSGOHEOASHGTVPBGHPSITWQOOSOPHGYVPPESVAAQGSOG 720
 Db 601 SSVFTESRVSQOQTVSYSGOHEOASHGTVPBGHPSITWQOOSOPHGYVPPESVAAQGSOG 660
 QY 721 QPSSSLLTGVSSSQPIQHPPOQOQGIQOTAPPOQTVYSLSQTSSTSEATTAQVPSOPAP 780
 Db 661 QPSSSLLTGVSSSQPIQHPPOQOQGIQOTAPPOQTVYSLSQTSSTSEATTAQVPSOPAP 720
 QY 781 QVLPVPSAGQO----- 791
 Db 721 QVLPVPSAGQO----- 791
 QY 791 QVLPVPSAGQO----- 791
 Db 781 VSOQIPSTPHVSTAQGTSSSLPTMAAGITOPILTLLASATVTAIPGVSTVVPBOLPTLL 840
 QY 791 ----- 791
 Db 841 QPVTOLPSQVHPOLLQPAYOSWGIIPANLGOAEVPLSSGDUVLYOGFPPLRPPOYPGDSNT 900
 QY 791 ----- 791
 Db 901 APSSNVAVCIHSTVLAPMPTEVLATPGYFPVTVQPYVESNLLVPMGAGVGGQVQVQSPG 960
 QY 791 ----- 791
 Db 961 GSILAQAPTSSQOAVLESTQGVSOVAPAEVVAQAOPATQPTLLASVDASHDVASGMS 1020
 QY 835 DGENEVSSSSGRHGRITTKHRYKSVASRHEKTSRPKLILINVSNGKGRVVECOLETH 894
 Db 1021 DGENEVSSSSGRHGRITTKHRYKSVASRHEKTSRPKLILINVSNGKGRVVECOLETH 894
 QY 895 NRGKMTFKPLDNDNPELITIMVNDFLAIERESVVDVREIIEKADMLSEDEVSEP 954
 Db 1081 NRGKMTFKPLDNDNPELITIMVNDFLAIERESVVDVREIIEKADMLSEDEVSEP 954
 QY 955 EGDQGESLQGDYDGFSGSOXLEGEFKOPIPASWPMOQIGPTSSLTQVYHAGRRFY 1014
 Db 1141 EGDQGESLQGDYDGFSGSOXLEGEFKOPIPASWPMOQIGPTSSLTQVYHAGRRFY 1014
 QY 1015 SPVPSERLRESKYVPESEITDTVAASTAOSPMMNLSHASSLSLQOAFSELRAQOMTEGPN 1074
 Db 1201 SPVPSERLRESKYVPESEITDTVAASTAOSPMMNLSHASSLSLQOAFSELRAQOMTEGPN 1074
 QY 1075 TAPPNSSHGTFTPPVVPPLFSSITAGVPTTAATAFVPAISSPPNDISTSVTQSEVTPTE 1134
 Db 1261 TAPPNSSHGTFTPPVVPPLFSSITAGVPTTAATAFVPAISSPPNDISTSVTQSEVTPTE 1134

QY 1135 EGIAGVNTSGVNTSGGLPIPPVSESPVSSVSSITIPAVVSIITSPSSQVPTSEI 1194
 Db 1321 EGIAGVNTSGVNTSGGLPIPPVSESPVSSVSSITIPAVVSIITSPSSQVPTSEI 1194
 QY 1195 VVSSTALYPSTVATASAGSTATPGKPPAVVSOQAASITVGATLTSVSTTSPS 1254
 Db 1381 VVSSTALYPSTVATASAGSTATPGKPPAVVSOQAASITVGATLTSVSTTSPS 1254
 QY 1255 TASQLSIOLSSSTPTLAEVTVVVAHSIDKTSHTSTGLAFSLAPSSSSPGAGVSY 1314
 Db 1441 TASQLSIOLSSSTPTLAEVTVVVAHSIDKTSHTSTGLAFSLAPSSSSPGAGVSY 1314
 QY 1315 ISOPGGLHPLVYPSVASTPILPQAGPTSPILPQVPSIPPLVQPVANVAVOQTLHS 1374
 Db 1501 ISOPGGLHPLVYPSVASTPILPQAGPTSPILPQVPSIPPLVQPVANVAVOQTLHS 1374
 QY 1375 QOPALLPNOPTHCPVDSDTOPKAPGIDDIKTLERKLSFSEHSSSGAQHVSLET 1434
 Db 1561 QOPALLPNOPTHCPVDSDTOPKAPGIDDIKTLERKLSFSEHSSSGAQHVSLET 1434
 QY 1435 SLVIESTVTGPIPTTAVAPSKLITSTSTCLPPTNLPLGTVALPVTVPVPGQVSTVST 1494
 Db 1621 SLVIESTVTGPIPTTAVAPSKLITSTSTCLPPTNLPLGTVALPVTVPVPGQVSTVST 1494
 QY 1495 TTSGVKRGTPASKPEPLTKAVLPGTELPAQTLPSBOLPPPPGSLTQSOQPLEDLDAOL 1554
 Db 1681 TTSGVKRGTPASKPEPLTKAVLPGTELPAQTLPSBOLPPPPGSLTQSOQPLEDLDAOL 1554
 QY 1555 RRTLSPEXITVTSVAVGVMAAPALTEAGTOPKQSOVKEGVLAATSGAGVFKGRF 1614
 Db 1741 RRTLSPEXITVTSVAVGVMAAPALTEAGTOPKQSOVKEGVLAATSGAGVFKGRF 1614
 QY 1615 QVSVADGAQKEGKNSEDAKSVHFESESSSVLSSSSBESTLYKEPNGITITPGISD 1674
 Db 1801 QVSVADGAQKEGKNSEDAKSVHFESESSSVLSSSSBESTLYKEPNGITITPGISD 1674
 QY 1675 VPESAHTTSAEAKSDTGOPTKVGPFQVTTTANKVGFVSUKTEDKTTDTKKGVPVAPSP 1734
 Db 1861 VPESAHTTSAEAKSDTGOPTKVGPFQVTTTANKVGFVSUKTEDKTTDTKKGVPVAPSP 1734
 QY 1735 FMDLEQAVLPVAVIPKKEKPELSEPSHNGSPSDEAFSLRDVDDGSGSPHSHOLSSKS 1794
 Db 1921 FMDLEQAVLPVAVIPKKEKPELSEPSHNGSPSDEAFSLRDVDDGSGSPHSHOLSSKS 1794
 QY 1795 LPSQNLSSQSLSNSFNSSSYMSDNESEDIEDLKLLELRRLDKLKEIQDLQSHQKHEIES 1854
 Db 1981 LPSQNLSSQSLSNSFNSSSYMSDNESEDIEDLKLLELRRLDKLKEIQDLQSHQKHEIES 1854
 QY 1855 LYTUKLGVPPVAVIIPPAAPLSGRRRPTYSKSGKSSRSSSLGKNSPOLSGNLSGQSAASV 1914
 Db 2041 LYTUKLGVPPVAVIIPPAAPLSGRRRPTYSKSGKSSRSSSLGKNSPOLSGNLSGQSAASV 1914
 QY 1915 LHPQOQLHPPGNIPESGQOQLQPLKPSPSDMLYGAFTSDGASVPSLSAPGQGSTSTN 1974
 Db 2101 LHPQOQLHPPGNIPESGQOQLQPLKPSPSDMLYGAFTSDGASVPSLSAPGQGSTSTN 1974
 QY 2101 LHPQOQLHPPGNIPESGQOQLQPLKPSPSDMLYGAFTSDGASVPSLSAPGQGSTSTN 2160
 Db 2161 LHPQOQLHPPGNIPESGQOQLQPLKPSPSDMLYGAFTSDGASVPSLSAPGQGSTSTN 2160
 QY 2035 MARKFSAPQOLCISMTSNIGGSAPIGAASATSGHTTTSKCPQOQGFPAITPGAGMSGT 2094
 Db 2221 MARKFSAPQOLCISMTSNIGGSAPIGAASATSGHTTTSKCPQOQGFPAITPGAGMSGT 2094
 QY 2095 GGPAPQPLQOPQVGTASLQNFNISMLQKSIINPPPSNLRRT 2136
 Db 2281 GGPAPQPLQOPQVGTASLQNFNISMLQKSIINPPPSNLRRT 2136

RESULT 10
 US-09-854-856-62
 ; Sequence 62, Application us/09854856

Patent No. 6541252
GENERAL INFORMATION:
APPLICANT: Walke, D. Wade
APPLICANT: Hilbun, Erin
APPLICANT: Donoho, Gregory
APPLICANT: Turner, C. Alexander Jr.
TITLE OF INVENTION: No. 6541252 Human Kinases and Polynucleotides
FILE REFERENCE: Encoding the Same
CURRENT APPLICATION NUMBER: US-09/854,856
CURRENT FILING DATE: 2001-05-14
PRIORITY APPLICATION NUMBER: US 60/206,015
PRIORITY FILING DATE: 2000-05-19
NUMBER OF SEQ ID NOS: 64
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 62
LENGTH: 2048
TYPE: PRN
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: VARIANT
LOCATION: (1)...(2048)
OTHER INFORMATION: Xaa = Any Amino Acid
US-09-854-856-62

Query Match 95.7%; Score 10344; DB 4; Length 2048;
Best Local Similarity 98.7%; Pred. No. 0;
Matches 2048; Conservative 0; Mismatches 0; Indels 28; Gaps 1;

61 MDKDSRGAATTTTTEHFRFRSVICDSNATLELPLSLPLPSIPAAPVPOGAPPEPH 120
1 MDKDSRGAATTTTTEHFRFRSVICDSNATLELPLSLPLPSIPAAPVPOGAPPEPH 60
121 REETVTTATSOVAQOQPPAAAAPGEOAVAGPAPSTVSSSKORPVOPSLVSGKEPP 180
61 REETVTTATSOVAQOQPPAAAAPGEOAVAGPAPSTVSSSKORPVOPSLVSGKEPP 120
181 ARSGSGGSAKEPEERSQOQDDIEELKAVGMSNGRFLKFDIEGRSGFKTVYGLD 240
121 ARSGSGGSAKEPEERSQOQDDIEELKAVGMSNGRFLKFDIEGRSGFKTVYGLD 180
241 TETVEVAMCELODRKLTKEERORFKEAEMLKGLQHPTVRYFSDWESTVKGKCIIVL 300
181 TETVEVAMCELODRKLTKEERORFKEAEMLKGLQHPTVRYFSDWESTVKGKCIIVL 240
301 TELMTSGTLKTYLKRFRVMKTVLRSCROILKGLQFLHRTPTPIHRDLKCDNIFITGP 360
241 TELMTSGTLKTYLKRFRVMKTVLRSCROILKGLQFLHRTPTPIHRDLKCDNIFITGP 300
361 TGSVKTIGLGLATLKRPASFAKSVIGTPEFMAPEMYEEKYDESVDVYAFGMCLMATSEY 420
301 TGSVKTIGLGLATLKRPASFAKSVIGTPEFMAPEMYEEKYDESVDVYAFGMCLMATSEY 360
421 PYSECQNAQAQYRREVTSGVCPASFDKVAIPEVKEIIIEGCIROKNDERYSIKDNLNHAFFQ 480
361 PYSECQNAQAQYRREVTSGVCPASFDKVAIPEVKEIIIEGCIROKNDERYSIKDNLNHAFFQ 420
481 BETGVRLAEEDDEGEKIAIKLMLRIEDIKLKGKIKYKONKNAIESFDLERVPEDVAQEM 540
421 BETGVRLAEEDDEGEKIAIKLMLRIEDIKLKGKIKYKONKNAIESFDLERVPEDVAQEM 480
541 VESGVGCGDHKTAKAIKDRVSLIKRREOROLVREOEKKOEKKESSLKQOVESASQ 600
481 VESGVGCGDHKTAKAIKDRVSLIKRREOROLVREOEKKOEKKESSLKQOVESASQ 540
601 TGIRQLPSASTGIPASTTSASVSTOVEPEEPADQOOLQYQOPSISVLSDGTVDSSQG 660
541 TGIRQLPSASTGIPASTTSASVSTOVEPEEPADQOOLQYQOPSISVLSDGTVDSSQG 600
661 SSVTFESVSSQOQTVSYSGHBOAHSTGTVEGHIPSTYQAOOSOPHGVPPSSVAGQSGQ 720
601 SSVTFESVSSQOQTVSYSGHBOAHSTGTVEGHIPSTYQAOOSOPHGVPPSSVAGQSGQ 660

721 QPSSSLTGVSSSQPIQHPQOQOQIGIOQTAPPOQTVYSLISQTSSEATTAPQVPOQAP 780
654 -----QQIGQITAPPOQTVYSLISQTSSEATTAPQVPOQAP 692
781 QVLPQVSAKQSTQGVSOVAPAEVVAOPATOPPTLASSVDSASHSDVAGSGDGENV 840
693 QVLPQVSAKQSTQGVSOVAPAEVVAOPATOPPTLASSVDSASHSDVAGSGDGENV 752
841 PSSSGHGRRTTKRHRYKSVRSRHEKTSRPLRLILANSNKDRAVECOLETHNRKMT 900
753 PSSSGHGRRTTKRHRYKSVRSRHEKTSRPLRLILANSNKDRAVECOLETHNRKMT 812
901 FKFDLDGNDPEIATIMVNDFTLAIERESFVDQVEIIEKADMLSEDSVVEPDQGL 960
813 FKFDLDGNDPEIATIMVNDFTLAIERESFVDQVEIIEKADMLSEDSVVEPDQGL 872
961 ESIQKDDYGFSSQKLEGEFKQPIASSMPQOIGIPTSLLQVHSAGRFTVSPVES 1020
873 ESIQKDDYGFSSQKLEGEFKQPIASSMPQOIGIPTSLLQVHSAGRFTVSPVES 932
1021 RLRESKVPSEITDTVAASTAQPQGNLSSHASLSLQCAFSELRRAQMTGENTAPPNF 1080
933 RLRESKVPSEITDTVAASTAQPQGNLSSHASLSLQCAFSELRRAQMTGENTAPPNF 992
1081 SHTGPTFPVVPPLSSIAQVPTTAAATAPVATSSPPNDISTSVTQSEVTPPEEGTAGV 1140
993 SHTGPTFPVVPPLSSIAQVPTTAAATAPVATSSPPNDISTSVTQSEVTPPEEGTAGV 1052
1141 ATSGVATSGGLPIPPVPSBPVLSVSSITTPAVVISITTSLSLOVPTSEIIVSSTA 1200
1053 ATSGVATSGGLPIPPVPSBPVLSVSSITTPAVVISITTSLSLOVPTSEIIVSSTA 1112
1201 LYPSTVATSASAGSTATPGPKPAPVAVVSOQAAGSTTVGATLTSVTTTSPSTASQLS 1260
1113 LYPSTVATSASAGSTATPGPKPAPVAVVSOQAAGSTTVGATLTSVTTTSPSTASQLS 1172
1261 IQLSSSTPTLAEVTVVASHSLDKTSHSSTTGLASLSAPSSSSSPGACVSYISOPGG 1320
1173 IQLSSSTPTLAEVTVVASHSLDKTSHSSTTGLASLSAPSSSSSPGACVSYISOPGG 1232
1321 LHPVLVPSVASTPILPQAAGPTSTPLLPVPSIPPLVQVAVVPAVQOQLIHSQOPAL 1380
1233 LHPVLVPSVASTPILPQAAGPTSTPLLPVPSIPPLVQVAVVPAVQOQLIHSQOPAL 1292
1381 LHPQPTHCEVNDSDTPKAPGIDDKITLBEKLSLSESHSSGQAQHASVLESTLVIES 1440
1293 LHPQPTHCEVNDSDTPKAPGIDDKITLBEKLSLSESHSSGQAQHASVLESTLVIES 1352
1441 TVTPGIPPTAVAPSKLLTSTSTCLPPTNLPLGTVALPVTVPVTPGOVSTPVSTTSGVK 1500
1353 TVTPGIPPTAVAPSKLLTSTSTCLPPTNLPLGTVALPVTVPVTPGOVSTPVSTTSGVK 1412
1501 PGRAPSRPPLTKAPVLVPGELPAGTLPSBOLPPFPGPSLTOGOQPLEDDAOLRRLTSP 1560
1413 PGRAPSRPPLTKAPVLVPGELPAGTLPSBOLPPFPGPSLTOGOQPLEDDAOLRRLTSP 1472
1561 EXITVTSVAVPVSAAATTAITEAGTQPOKGVSYQYKEGPVLTASSGACVFKMGRPOVAA 1620
1473 EXITVTSVAVPVSAAATTAITEAGTQPOKGVSYQYKEGPVLTASSGACVFKMGRPOVAA 1532
1621 DGAQKEGKNSKEDAKSVHESSTESSVSSSPESSTLVKPEPNTIIPGISVVPESA 1680
1533 DGAQKEGKNSKEDAKSVHESSTESSVSSSPESSTLVKPEPNTIIPGISVVPESA 1592
1681 KTTASEAKSDTQGTQYGRFOVTTTAKVGFVSXKEDKTDTPKKGGPVAPPPFMLEQ 1740
1593 KTTASEAKSDTQGTQYGRFOVTTTAKVGFVSXKEDKTDTPKKGGPVAPPPFMLEQ 1652
1741 AVLPVAVIPKKEKPELSEPHLNGSPSDEPAFLSRVDVDSGSPHSBHOLSKSLPSQNT 1800
1653 AVLPVAVIPKKEKPELSEPHLNGSPSDEPAFLSRVDVDSGSPHSBHOLSKSLPSQNT 1712
1801 SGLSNSFNSSYKMSDNESEDIEDDLKLELRRLDKLKEIQDLQSKQKHIEBLVYTKG 1860

Db 1713 SOSLSNSFNSSWSSSDNEDIEDDLKLELRRLRDKHKEIODLQSRQKHEIESLYTKLG 1772
 QY 1861 KVPAPVILPPAPLPSGRRRRPTKSGKSSRSSSLGNKSPOLSGSLGSOASAASVHPQOT 1920
 Db 1773 KVPAPVILPPAPLPSGRRRRPTKSGKSSRSSSLGNKSPOLSGSLGSOASAASVHPQOT 1832
 QY 1921 LHPNITRESGONQLQPLKPSRSSDNLVSAFTSDGALSVPSLSAPGQGTSTNTVGATV 1980
 Db 1833 LHPNITRESGONQLQPLKPSRSSDNLVSAFTSDGALSVPSLSAPGQGTSTNTVGATV 1892
 QY 1981 NSOAAQAPPMATSSRKGTFTDHLKLVDMNARDAMNSGRGSGKHNYEGPGMARFES 2040
 Db 1893 NSOAAQAPPMATSSRKGTFTDHLKLVDMNARDAMNSGRGSGKHNYEGPGMARFES 1952
 QY 2041 APGOLCISMTSNLGSAPISAAATSLGHFTKSMCPQOYGPAPTPGAGMSGTGGAPAPQ 2100
 Db 1953 APGOLCISMTSNLGSAPISAAATSLGHFTKSMCPQOYGPAPTPGAGMSGTGGAPAPQ 2012
 QY 2101 PLGQFQPVGTASLQNFNINLQKSIINPQSNLRAT 2136
 Db 2013 PLGQFQPVGTASLQNFNINLQKSIINPQSNLRAT 2048

RESULT 11

US-09-854-856-56
 / Sequence 56, Application US/09854856
 / Patent No. 6541252
 / GENERAL INFORMATION:
 / APPLICANT: Walke, D. Wade
 / APPLICANT: Hilbun, Erin
 / APPLICANT: Donoho, Gregory
 / APPLICANT: Turner, C. Alexander Jr.
 / TITLE OF INVENTION: No. 6541252el Human Kinases and Polynucleotides
 / FILE REFERENCE: LEX-0178-USA
 / CURRENT APPLICATION NUMBER: US/09/854,856
 / PRIOR FILING DATE: 2001-05-14
 / PRIOR APPLICATION NUMBER: US 60/206,015
 / NUMBER OF SEQ ID NOS: 64
 / SOFTWARE: FastSeq for Windows Version 4.0
 / SEQ ID NO 56
 / LENGTH: 2141
 / TYPE: PRT
 / ORGANISM: Homo sapiens
 / FEATURE:
 / NAME/KEY: VARIANT
 / LOCATION: (1)...(2141)
 / OTHER INFORMATION: Xaa = Any Amino Acid
 US-09-854-856-56

Query Match 95.1%; Score 10287.5; DB: 4; Length 2141;
 Best Local Similarity 94.4%; Pred. No. 0;
 Matches 2048; Conservative 0; Mismatches 0; Indels 121; Gaps 2;
 QY 61 MDKDSGAAAATTTTTHRRFRFRSVICDSNATALELPLSLPQPSIPAAVPOSAPEPEH 120
 Db 1 MDKDSGAAAATTTTTHRRFRFRSVICDSNATALELPLSLPQPSIPAAVPOSAPEPEH 60
 QY 121 RETTVATATSVQAQPPAAAPGEQAVGAPAPSTVPSSTSKORPVSOPLVSGKEEPP 180
 Db 61 RETTVATATSVQAQPPAAAPGEQAVGAPAPSTVPSSTSKORPVSOPLVSGKEEPP 120
 QY 181 AASGSGSAAKEQERSQOQDDIELETKAVMSNDGRFLKEDIEIGSGFEYVKKGLD 240
 Db 121 AASGSGSAAKEQERSQOQDDIELETKAVMSNDGRFLKEDIEIGSGFEYVKKGLD 180
 QY 241 TETTVAVANCELQDRKLTKSERQRFKEAEMLKGLQHPNIVRFYDSWESTVAKKCIIV 300
 Db 181 TETTVAVANCELQDRKLTKSERQRFKEAEMLKGLQHPNIVRFYDSWESTVAKKCIIV 240
 QY 301 TELMNSGLTKYLKRFKWKIKVLRSWCQILKGLQFLHTRTPIIHRDLKCDNIFITGP 360

Db 241 TELMNSGLTKYLKRFKWKIKVLRSWCQILKGLQFLHTRTPIIHRDLKCDNIFITGP 300
 QY 361 TGSYKIGDLGATIKRASFAKSVIGPEFMAPEWEEKCYDSVDYAGMCMLEMAATSEY 420
 Db 301 TGSYKIGDLGATIKRASFAKSVIGPEFMAPEWEEKCYDSVDYAGMCMLEMAATSEY 360
 QY 421 PYSCONAAQIYRVTSGVKASFDKVAIPVKEIIECCIFONKDERSTYDOLNHAFFQ 480
 Db 361 PYSCONAAQIYRVTSGVKASFDKVAIPVKEIIECCIFONKDERSTYDOLNHAFFQ 420
 QY 481 EETGVARELAEDDEGEKIAIKMLRIEDIKLKGYKQNEAIEFSFDERDVPENVAQEM 540
 Db 421 EETGVARELAEDDEGEKIAIKMLRIEDIKLKGYKQNEAIEFSFDERDVPENVAQEM 480
 QY 541 VESGYVEBGDKHTWAKAIKQVSLIKRRREGQRLVREOEKKOEESLKQOVBOSASQ 600
 Db 481 VESGYVEBGDKHTWAKAIKQVSLIKRRREGQRLVREOEKKOEESLKQOVBOSASQ 540
 QY 601 TGIKQLPASATGIPPTASVSTOVEPEPEADQHQLOQOQPSISVSDGTVDGQG 660
 Db 541 TGIKQLPASATGIPPTASVSTOVEPEPEADQHQLOQOQPSISVSDGTVDGQG 600
 QY 661 SSYFTESRVSSQQTVSYSQHEQAHSTGTVFGHLPSTVOAOSQPHGYPPSSVAQGSQ 720
 Db 601 SSYFTESRVSSQQTVSYSQHEQAHSTGTVFGHLPSTVOAOSQPHGYPPSSVAQGSQ 660
 QY 721 QPSSSLITGVSSSQPIQHPQOQCGIQQTPAPQOQVYSLQSTSSBATTAAQVPQOPAP 780
 Db 654 QPSSSLITGVSSSQPIQHPQOQCGIQQTPAPQOQVYSLQSTSSBATTAAQVPQOPAP 720
 QY 781 QVLPOVSAGKQPPPLPPQYPPGDSNAPSSNVAVCITHSVLXPPMPTVLAATPGYPT 752
 Db 693 QVLPOVSAGKQPPPLPPQYPPGDSNAPSSNVAVCITHSVLXPPMPTVLAATPGYPT 752
 QY 792 VQPYVESNLLVPMGGVGVGVQVVSQPGSLAQAPPTSSQQAIVLESTQGVQVAPAEVAV 807
 Db 753 VQPYVESNLLVPMGGVGVGVQVVSQPGSLAQAPPTSSQQAIVLESTQGVQVAPAEVAV 812
 QY 808 AQPQATQPTTLASVDSASDVASGMSDGENVPSSSGRHEGRTTKHYKVSRSRHE 867
 Db 813 AQPQATQPTTLASVDSASDVASGMSDGENVPSSSGRHEGRTTKHYKVSRSRHE 807
 QY 868 KTSRPKRLITANVKNKDRVVECOLETHNRKMTVEKPDLDGNDNEELATIMVNDPILATE 927
 Db 873 KTSRPKRLITANVKNKDRVVECOLETHNRKMTVEKPDLDGNDNEELATIMVNDPILATE 927
 QY 928 RESFVQVREIIEKADMLSEDSVVEPEGDQLESLQKQDYGFSGSOKLEGEFKQPIPA 987
 Db 933 RESFVQVREIIEKADMLSEDSVVEPEGDQLESLQKQDYGFSGSOKLEGEFKQPIPA 928
 QY 988 SSMPOQIGIPTSSLITOVVHSAGRFTVSPVPSRLRESKYFPESEITDTVAASQSPGMN 1047
 Db 993 SSMPOQIGIPTSSLITOVVHSAGRFTVSPVPSRLRESKYFPESEITDTVAASQSPGMN 1052
 QY 1048 LSHSASSLSIQQFSLRBAQMTGEPNTAPPNESHGPTFPVVPPLSLASAGPTTAAAT 1107
 Db 1053 LSHSASSLSIQQFSLRBAQMTGEPNTAPPNESHGPTFPVVPPLSLASAGPTTAAAT 1112
 QY 1108 APVPATSPENDISTSVIOSEVTVPTREGIAGVATGVTGSGGLPIPEVSESPVLSVV 1167
 Db 1113 APVPATSPENDISTSVIOSEVTVPTREGIAGVATGVTGSGGLPIPEVSESPVLSVV 1172
 QY 1168 SSTTPAVVISITTSLSLOVPTSTSIIVSSTNLVSVVATASAGGSTATPGKPPA 1227
 Db 1173 SSTTPAVVISITTSLSLOVPTSTSIIVSSTNLVSVVATASAGGSTATPGKPPA 1232
 QY 1228 VVSQQAAGSTYGATLTSVSTTSFPTASQSLQSSSTSTPTEAETVVAHSHDKDS 1287
 Db 1233 VVSQQAAGSTYGATLTSVSTTSFPTASQSLQSSSTSTPTEAETVVAHSHDKDS 1292
 QY 1288 HSGTGLAFLSLAPSSSSPGAGVSVYSQPGLHPLVPSVIASTPIIPQAGPSTPL 1347

Db 1293 HSSSTGLAFSLSPSSSSSGAGVSSYISQPGGLHPLVBSVIASTPILFQAAGPTSTPL 1352
 QY 1348 LPQVPSIPLVOPANVPANVOQTLIHSGQPALLPNOPTHCEPDSDTQPKAGIDIK 1407
 Db 1353 LPQVPSIPLVOPANVPANVOQTLIHSGQPALLPNOPTHCEPDSDTQPKAGIDIK 1412
 QY 1408 TLEBKSLFSEHSSSGAQAASVLETSVLESTVGTGIPITTAAPSKLLTSTSTCLPP 1467
 Db 1413 TLEBKSLFSEHSSSGAQAASVLETSVLESTVGTGIPITTAAPSKLLTSTSTCLPP 1472
 QY 1468 TNLPLGVALPTVPVPTPGQVSTVSTTSGVKRGTAAPSKPLTAVLPGTLEPAGTL 1527
 Db 1473 TNLPLGVALPTVPVPTPGQVSTVSTTSGVKRGTAAPSKPLTAVLPGTLEPAGTL 1532
 QY 1528 PSEQLPPEPGLSLQSQOPLDADOLRRTLSPEXITVTSVGVSAAPPAITAEATOP 1587
 Db 1533 PSEQLPPEPGLSLQSQOPLDADOLRRTLSPEXITVTSVGVSAAPPAITAEATOP 1592
 QY 1588 QKGVSYQKEGVLATSSGAGVFKMGRFOVSAADGAQKEGKNSKSDAKSVHFEESTSESS 1647
 Db 1593 QKGVSYQKEGVLATSSGAGVFKMGRFOVSAADGAQKEGKNSKSDAKSVHFEESTSESS 1652
 QY 1648 VLSSSESTLVKPEPNGITPGISSDVPEBAHKTASAKSDTQPKAGFQVITTTAN 1707
 Db 1653 VLSSSESTLVKPEPNGITPGISSDVPEBAHKTASAKSDTQPKAGFQVITTTAN 1712
 QY 1708 KVGFSYKTEKEDITDTRKKEGVPASPEFMDLEQAVLPAVLPKKEPELSEPSHLNGSPSD 1767
 Db 1713 KVGFSYKTEKEDITDTRKKEGVPASPEFMDLEQAVLPAVLPKKEPELSEPSHLNGSPSD 1772
 QY 1768 PEAFALSRDVEDGSGSPHSPHQLSSKSLPSQNLSSQSLNSFNSSYMSDNESEDIEDEDLK 1827
 Db 1773 PEAFALSRDVEDGSGSPHSPHQLSSKSLPSQNLSSQSLNSFNSSYMSDNESEDIEDEDLK 1832
 QY 1828 LELRLRLDKHLKEITODLOSROKHEIESLYTKLGKVPYPAVILPPAAPSGRRRPTKSKGS 1887
 Db 1833 LELRLRLDKHLKEITODLOSROKHEIESLYTKLGKVPYPAVILPPAAPSGRRRPTKSKGS 1892
 QY 1888 KSSRSSSTGKNSPOLSGMLSGQSAASVLAHPQOTLHPENIPESONOLQPLKSPSSDN 1947
 Db 1893 KSSRSSSTGKNSPOLSGMLSGQSAASVLAHPQOTLHPENIPESONOLQPLKSPSSDN 1952
 QY 1948 LYSAFSTDGAISSVLSLAPAGQSTSTNTVGAIVNSQAQAOPPAMTSSRKGTFTDIDLK 2007
 Db 1953 LYSAFSTDGAISSVLSLAPAGQSTSTNTVGAIVNSQAQAOPPAMTSSRKGTFTDIDLK 2012
 QY 2008 VDNARADAMNLSGRGSGKGMNRYBPGMARFSAPOGLCTIMTSLGSSAPISASATSL 2067
 Db 2013 VDNARADAMNLSGRGSGKGMNRYBPGMARFSAPOGLCTIMTSLGSSAPISASATSL 2072
 QY 2068 GHFTKSWCPPOOYGPATPFGAOWSGTGPAPOPFGQFQPGVGTASLQFNISNLOKSTSN 2127
 Db 2073 GHFTKSWCPPOOYGPATPFGAOWSGTGPAPOPFGQFQPGVGTASLQFNISNLOKSTSN 2132
 QY 2128 PPGSNLRTT 2136
 Db 2133 PPGSNLRTT 2141

; PRIOR APPLICATION NUMBER: US 60/206,015
 ; PRIOR FILING DATE: 2000-05-19
 ; NUMBER OF SEQ ID NOS: 64
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 50
 ; LENGTH: 2294
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: VARIANT
 ; LOCATION: (1)-(2294)
 ; OTHER INFORMATION: Xaa = Any Amino Acid
 ; US-09-854-856-50
 Query Match 94.4%; Score 10211; DB 4; Length 2294;
 Best Local Similarity 88.2%; Pred. No. 0;
 Matches 2048; Conservative 0; Mismatches 0; Indels 274; Gaps 2;
 Db 1 MDKDSRGAATTTTTEHFRFRSVCDSNATALEPGLPLSLPQPSIPAAVPOSAPPEPH 60
 QY 61 MDKDSRGAATTTTTEHFRFRSVCDSNATALEPGLPLSLPQPSIPAAVPOSAPPEPH 120
 Db 121 REETVATATSOVAQOPPAAPGEQAVAPASTVSSSKDRPVSOPLVSGKEEPP 180
 QY 121 REETVATATSOVAQOPPAAPGEQAVAPASTVSSSKDRPVSOPLVSGKEEPP 180
 Db 61 REETVATATSOVAQOPPAAPGEQAVAPASTVSSSKDRPVSOPLVSGKEEPP 120
 QY 181 ARSGSGGSAKEQOEESQOQDDIELETRAVGMSNDGRFLKPDIEIGRSFKTVXGLD 240
 Db 121 ARSGSGGSAKEQOEESQOQDDIELETRAVGMSNDGRFLKPDIEIGRSFKTVXGLD 180
 QY 241 TETTVAVMCELQDRKLTKEERORFKEAEMLKGLQHPNIVRYFDSWESTYKGCCTLV 300
 Db 181 TETTVAVMCELQDRKLTKEERORFKEAEMLKGLQHPNIVRYFDSWESTYKGCCTLV 240
 QY 301 TELMTSGTLKTYLKRFKVMKIVLRSMCRQILKGLQHLHRTPTIHRDLKCDNIPTGP 360
 Db 241 TELMTSGTLKTYLKRFKVMKIVLRSMCRQILKGLQHLHRTPTIHRDLKCDNIPTGP 300
 QY 361 TGSVKIDGLATLKRAFPKSVIGTEPFNAPEMYEKEKYESVDVYAFGMCMLEMATSEY 420
 Db 301 TGSVKIDGLATLKRAFPKSVIGTEPFNAPEMYEKEKYESVDVYAFGMCMLEMATSEY 360
 QY 421 PYSECQNAQOYRRVTSGVKPAFDKVAIPBEKELIGCTIRONDERYSIKDLNHAFFQ 480
 Db 361 PYSECQNAQOYRRVTSGVKPAFDKVAIPBEKELIGCTIRONDERYSIKDLNHAFFQ 420
 QY 481 EETGVRYELAEEDGEEKIALKMLRIDIKLKXKYDNEAIEFSFDERVPEPDADEM 540
 Db 421 EETGVRYELAEEDGEEKIALKMLRIDIKLKXKYDNEAIEFSFDERVPEPDADEM 480
 QY 541 VESGYVEGDHKTAKAIKDRVSLIKRREORQLVREOEKKQOESSLKQOVSASQ 600
 Db 481 VESGYVEGDHKTAKAIKDRVSLIKRREORQLVREOEKKQOESSLKQOVSASQ 540
 QY 601 TGIRKOLBSASTGIPASTTASVSTOYVEPEEPADQHOQLOLOQOQPSISVLSGTVDSGQ 660
 Db 541 TGIRKOLBSASTGIPASTTASVSTOYVEPEEPADQHOQLOLOQOQPSISVLSGTVDSGQ 600
 QY 661 SSVFTESSVSSQOQVTSVGSQHEQASHSTGVGHPSTYQAOOSOPHYVPSSVAQGSQ 720
 Db 601 SSVFTESSVSSQOQVTSVGSQHEQASHSTGVGHPSTYQAOOSOPHYVPSSVAQGSQ 660
 QY 721 QPSSSLLTGVSQOPQHPOQOQIGQTPAPQOQVYSLQSTSSSEATTAQPVSQPQAF 780
 Db 661 QPSSSLLTGVSQOPQHPOQOQIGQTPAPQOQVYSLQSTSSSEATTAQPVSQPQAF 720
 QY 781 QVLQVSAKQO-----QGGIGQAPQOQVYSLQSTSSSEATTAQPVSQPQAF 791
 Db 693 QVLQVSAKQO-----QGGIGQAPQOQVYSLQSTSSSEATTAQPVSQPQAF 752
 QY 792 ----- 791
 Db 753 VSOIPISTPHVSTAGTGFSSLPITMAAGITOPDLTLASSATTAAIPGVSTVPSQLPILL 812

QY 792 ----- 791
 Db 813 QPVTQLPSQVHQLLPAPVQSMGIPANIGQAAEVLSSGDVLVYGFPPLRPPEQVPGDNT 872
 QY 792 ----- 791
 Db 873 APSSNVA5C1HSITVLXPMPTEVLATPGYPTVVQPYVESNLVPMWGQGVQVQVSGP 932
 QY 792 ----- 791
 Db 933 GSIAOAPRTSSQOAVLSTQGV5QVAPAEPAVAQPOATQPTTLASSVDSASHSDVASCMS 992
 QY 835 DGENENVPSSSGRHEGRTTKRHRYKRSVR5RSRHEKTSRKLRLILN5NKGDAVVEQOLETH 894
 Db 993 DGENENVPSSSGRHEGRTTKRHRYKRSVR5RSRHEKTSRKLRLILN5NKGDAVVEQOLETH 1052
 QY 895 NRKAVTFKFDLDGDNPEEIIATIMVNDPILAIERESFVDQVREIIEKADEMI5EDV5VEP 954
 Db 1053 NRKAVTFKFDLDGDNPEEIIATIMVNDPILAIERESFVDQVREIIEKADEMI5EDV5VEP 1112
 QY 955 EGDGLESLOGKDYGFSGSQKLEEFKQPIPASGMPQOIGIPSSSLTQVYH5AGRRITV 1014
 Db 1113 EGDGLESLOGKDYGFSGSQKLEEFKQPIPASGMPQOIGIPSSSLTQVYH5AGRRITV 1172
 QY 1015 SPVESRLRESKVPSEITTDVAA5TAQSPGMNLSH5ASLSLQOAFSELRRAQMTGEPN 1074
 Db 1173 SPVESRLRESKVPSEITTDVAA5TAQSPGMNLSH5ASLSLQOAFSELRRAQMTGEPN 1232
 QY 1075 TAPNFSHTGTFPVVPPFLSSINGVPTTAATAAPVPA5SSPPNDISTSVI5QSEVTVPE 1134
 Db 1233 TAPNFSHTGTFPVVPPFLSSINGVPTTAATAAPVPA5SSPPNDISTSVI5QSEVTVPE 1292
 QY 1135 EGIAGVASTSTVNV5GGLPIPV5ESPL5SV5SITIPAV5IST5SP5I5QVTS5E1 1194
 Db 1293 EGIAGVASTSTVNV5GGLPIPV5ESPL5SV5SITIPAV5IST5SP5I5QVTS5E1 1352
 QY 1195 VVS5TALPV5TV5ATS5ASAG5TATPGKPPAV55QOAA5STVGA5TL5SV5T5T5FP5 1254
 Db 1353 VVS5TALPV5TV5ATS5ASAG5TATPGKPPAV55QOAA5STVGA5TL5SV5T5T5FP5 1412
 QY 1255 TAOGL5QL5SS5T5PT5LA5TV5V5AS5LDT5H5S5T5T5G5AF5S5AP5SS5SP5AG5V5S 1314
 Db 1413 TAOGL5QL5SS5T5PT5LA5TV5V5AS5LDT5H5S5T5T5G5AF5S5AP5SS5SP5AG5V5S 1472
 QY 1315 ISO5GG5H5PI5V5P5IA5ST5PL5PO5AG5T5PL5PO5PS5IP5LV5Q5V5AN5P5AV5Q5TL5H5 1374
 Db 1473 ISO5GG5H5PI5V5P5IA5ST5PL5PO5AG5T5PL5PO5PS5IP5LV5Q5V5AN5P5AV5Q5TL5H5 1532
 QY 1375 QOP5ALL5P5N5P5HT5CP5E5V5D5T5OP5K5P5ID5IK5LE5K5LS5F5EH5SS5GA5Q5H5AS5LET 1434
 Db 1533 QOP5ALL5P5N5P5HT5CP5E5V5D5T5OP5K5P5ID5IK5LE5K5LS5F5EH5SS5GA5Q5H5AS5LET 1592
 QY 1435 SLV5IES5TV5TG5IPT5TA5V5AS5KL5T5T5T5T5CL5P5P5NL5PL5G5T5V5AL5P5TV5V5T5PG5V5T5EV5T 1494
 Db 1593 SLV5IES5TV5TG5IPT5TA5V5AS5KL5T5T5T5T5CL5P5P5NL5PL5G5T5V5AL5P5TV5V5T5PG5V5T5EV5T 1652
 QY 1495 TTS5GV5K5GT5AP5RP5L5TK5AP5VL5P5V5TE5L5P5AG5TL5PS5BD5LP5FP5PS5L5T5Q5O5P5L5ED5DA5Q 1554
 Db 1653 TTS5GV5K5GT5AP5RP5L5TK5AP5VL5P5V5TE5L5P5AG5TL5PS5BD5LP5FP5PS5L5T5Q5O5P5L5ED5DA5Q 1712
 QY 1555 RRT5L5PEX5IT5V5T5AV5GV5MA5P5AL5TE5AG5T5OP5K5G5V5Q5KE5GP5VL5AT5SG5AG5V5FK5M5GR5F 1614
 Db 1713 RRT5L5PEX5IT5V5T5AV5GV5MA5P5AL5TE5AG5T5OP5K5G5V5Q5KE5GP5VL5AT5SG5AG5V5FK5M5GR5F 1772
 QY 1615 QVS5V5AAD5G5AQ5EK5K5N5S5ED5K5SV5H5F5S5T5S5ESS5L5SS5S5P5E5ST5LV5PE5P5G5IT5PG5ISS5D 1674
 Db 1773 QVS5V5AAD5G5AQ5EK5K5N5S5ED5K5SV5H5F5S5T5S5ESS5L5SS5S5P5E5ST5LV5PE5P5G5IT5PG5ISS5D 1832
 QY 1675 VP5ESA5K5TT5A5S5E5K5ST5G5OPT5K5V5GR5FOV5TT5ANK5V5GR5FS5K5ED5IT5D5PK5KE5GP5V5A5SP 1734
 Db 1833 VP5ESA5K5TT5A5S5E5K5ST5G5OPT5K5V5GR5FOV5TT5ANK5V5GR5FS5K5ED5IT5D5PK5KE5GP5V5A5SP 1892

QY 1735 FMDLEQAVLPAVTPKKEPELSEPSHLN5G5SD5PEAAFL5SDV5D5G5SG5PH5S5PHO5LS5K5 1794
 Db 1893 FMDLEQAVLPAVTPKKEPELSEPSHLN5G5SD5PEAAFL5SDV5D5G5SG5PH5S5PHO5LS5K5 1952
 QY 1795 LPSQNL5Q5LS5NS5FN5SY5MS5DN5SD5IED5DL5K5EL5RL5R5KH5LE5IOD5L5OR5QH5IES 1854
 Db 1953 LPSQNL5Q5LS5NS5FN5SY5MS5DN5SD5IED5DL5K5EL5RL5R5KH5LE5IOD5L5OR5QH5IES 2012
 QY 1855 LYTL5GK5V5PAV5IIP5PA5L5SG5RR5R5PT5K5SG5K5SR5SS5LG5N5K5P5Q5L5SG5SA5V 1914
 Db 2013 LYTL5GK5V5PAV5IIP5PA5L5SG5RR5R5PT5K5SG5K5SR5SS5LG5N5K5P5Q5L5SG5SA5V 2072
 QY 1915 LHPQOT5LHP5GN5IES5Q5N5Q5L5Q5L5K5P5SS5DN5I5S5AT5SG5ALS5V5P5LS5AP5G5T5STN 1974
 Db 2073 LHPQOT5LHP5GN5IES5Q5N5Q5L5Q5L5K5P5SS5DN5I5S5AT5SG5ALS5V5P5LS5AP5G5T5STN 2132
 QY 1975 TVGATVNS5QAA5QAP5P5MT5SR5K5GT5FT5DD5L5K5L5VD5N5AR5D5AM5N5L5SG5RG5K5GH5N5Y5B5G 2034
 Db 2133 TVGATVNS5QAA5QAP5P5MT5SR5K5GT5FT5DD5L5K5L5VD5N5AR5D5AM5N5L5SG5RG5K5GH5N5Y5B5G 2192
 QY 2035 MARK5SAP5QOL5C5MT5SN5L5G5S5AP5ISA5S5AT5SL5GH5FT5K5M5CP5Q5Y5GP5AT5P5GA5W5SGT 2094
 Db 2193 MARK5SAP5QOL5C5MT5SN5L5G5S5AP5ISA5S5AT5SL5GH5FT5K5M5CP5Q5Y5GP5AT5P5GA5W5SGT 2252
 QY 2095 GGPAP5P5L5G5P5Q5P5V5GT5AS5L5Q5N5FN5SN5L5OK5S5N5P5GS5N5L5RT 2136
 Db 2253 GGPAP5P5L5G5P5Q5P5V5GT5AS5L5Q5N5FN5SN5L5OK5S5N5P5GS5N5L5RT 2294

RESULT 13
 US-09-854-856-16
 : Sequence 16, Application US/09854856
 : Patent No. 6541252
 : GENERAL INFORMATION:
 : APPLICANT: Walke, D. Wade
 : APPLICANT: Hilbun, Erin
 : APPLICANT: Donoho, Gregory
 : APPLICANT: Turner, C. Alexander Jr.
 : TITLE OF INVENTION: No. 6541252el Human Kinases and Polynucleotides
 : TITLE OF INVENTION: Encoding the Same
 : FILE REFERENCE: LEX-0178-USA
 : CURRENT APPLICATION NUMBER: US/09/854,856
 : CURRENT FILING DATE: 2001-05-14
 : PRIOR APPLICATION NUMBER: US 60/206,015
 : PRIOR FILING DATE: 2000-05-19
 : NUMBER OF SEQ ID NOS: 64
 : SOFTWARE: FastSeq for Windows Version 4.0
 : SEQ ID NO 16
 : LENGTH: 1999
 : TYPE: PRT
 : ORGANISM: Homo sapiens
 : FEATURE:
 : NAME/KEY: VARIANT
 : LOCATION: (1)...(1999)
 : OTHER INFORMATION: Xaa = Any Amino Acid
 US-09-854-856-16

Query Match 91.8%; Score 9929; DB 4; Length 1999;
 Best Local Similarity 99.7%; Pred. No. 0;
 Matches 1972; Conservative 1; Mismatches 1; Indels 4; Gaps 1;

QY 1 MSGGA5K5Q5ST5P5G5L5L5P5P5AP5K5N5G5S5D5S5V5G5K5L5G5AA5A5D5AV5T5G5T5E5Y5R5R5R5T 60
 Db 1 MSGGA5K5Q5ST5P5G5L5L5P5P5AP5K5N5G5S5D5S5V5G5K5L5G5AA5A5D5AV5T5G5T5E5Y5R5R5R5T 60
 QY 61 MDKDS5GA5A5T5T5T5T5H5R5F5R5S5V5IC5D5N5AT5ALE5L5P5L5P5Q5S5IP5A5V5P5Q5A5P5E5P5H 120
 Db 61 MDKDS5GA5A5T5T5T5T5H5R5F5R5S5V5IC5D5N5AT5ALE5L5P5L5P5Q5S5IP5A5V5P5Q5A5P5E5P5H 120
 QY 121 REETV5AT5ATS5Q5V5AQ5P5P5AA5P5G5Q5AV5AG5P5AB5T5V5S5T5SK5OR5PV5Q5S5LV5G5KE5E5P5P 180
 Db 121 REETV5AT5ATS5Q5V5AQ5P5P5AA5P5G5Q5AV5AG5P5AB5T5V5S5T5SK5OR5PV5Q5S5LV5G5KE5E5P5P 180

QY 181 ARSGSGGSAKPEEERSQOQDDIELEETAYAGKMSNDGRFLKEDIEIGRGSFKTVYKGLD 240
 Db 181 ARSGSGGSAKPEEERSQOQDDIELEETAYAGKMSNDGRFLKEDIEIGRGSFKTVYKGLD 240
 QY 241 TETTVAVANGELQDRKLTSEBQRFKEAEMLKGLQHPNIVRFPDSESTVYKGCCLVY 300
 Db 241 TETTVAVANGELQDRKLTSEBQRFKEAEMLKGLQHPNIVRFPDSESTVYKGCCLVY 300
 QY 301 TELMTSGTLKTYLKRFFVMKIKVLRSCROILKGLQHLHTRTPPIIHRDKCDNIFITGP 360
 Db 301 TELMTSGTLKTYLKRFFVMKIKVLRSCROILKGLQHLHTRTPPIIHRDKCDNIFITGP 360
 QY 361 TGSVKIGDLGLATLKRAFAKSVIGTEPEMAPEYKEKDESVVDVAFGCMLEMATSEY 420
 Db 361 TGSVKIGDLGLATLKRAFAKSVIGTEPEMAPEYKEKDESVVDVAFGCMLEMATSEY 420
 QY 421 PYBRCQNAAQIYRVTSGVVPASFDKVAIIBVKEIIEGCTRONDEKYSIKDLINAFQ 480
 Db 421 PYBRCQNAAQIYRVTSGVVPASFDKVAIIBVKEIIEGCTRONDEKYSIKDLINAFQ 480
 QY 481 BETGVRELAEDEGEKIAIKMLRIDIKKLKGYKONEAIEFSPLEBDVPEDAQEM 540
 Db 481 BETGVRELAEDEGEKIAIKMLRIDIKKLKGYKONEAIEFSPLEBDVPEDAQEM 540
 QY 541 VESGYCEGDHKTAKAIKORVSLIKRRBQROLVREEQEKKESSLMQOVBQSSAQ 600
 Db 541 VESGYCEGDHKTAKAIKORVSLIKRRBQROLVREEQEKKESSLMQOVBQSSAQ 600
 QY 601 TGIQOLPSASTGIFTASTSASTOYEPPEPADQOQLOQOOPSISVSLDGTVDGQG 660
 Db 601 TGIQOLPSASTGIFTASTSASTOYEPPEPADQOQLOQOOPSISVSLDGTVDGQG 660
 QY 661 SSVFTESSVSSQOQVVSQSGOHEOASHGTGVGHIPSTVOAOSOPHGYPPSSVAQOSQ 720
 Db 661 SSVFTESSVSSQOQVVSQSGOHEOASHGTGVGHIPSTVOAOSOPHGYPPSSVAQOSQ 720
 QY 721 QPSSSILTVSSSQPIQHPQOQGIQOTAPPOQTVOYSLSQSTSSSEATTAQVPSQAP 780
 Db 721 QPSSSILTVSSSQPIQHPQOQGIQOTAPPOQTVOYSLSQSTSSSEATTAQVPSQAP 780
 QY 781 QVLPQVSAKOSTQGVSOVAPABEVAAQOPATOPTLLASVDSASHSVASGMSDGENV 840
 Db 781 QVLPQVSAKOSTQGVSOVAPABEVAAQOPATOPTLLASVDSASHSVASGMSDGENV 840
 QY 841 PSSSGRHEGRTTKHRYKSVASRHEKTSRPKRLINVSINKGRVYECOLETNRKMYT 900
 Db 841 PSSSGRHEGRTTKHRYKSVASRHEKTSRPKRLINVSINKGRVYECOLETNRKMYT 900
 QY 901 FKFDLDGDNPEELATIMVNDPILAIERESFVDQVREIIEKADMLSEDSVEBEDQGL 960
 Db 901 FKFDLDGDNPEELATIMVNDPILAIERESFVDQVREIIEKADMLSEDSVEBEDQGL 960
 QY 961 ESLOGKDDYGRSSGQKLEGEFKQPIPASSMPOQIGIPSSSLTVYVHSGRFFIYSPVES 1020
 Db 961 ESLOGKDDYGRSSGQKLEGEFKQPIPASSMPOQIGIPSSSLTVYVHSGRFFIYSPVES 1020
 QY 1021 RLRSKVPSEITITVAASTQSPQMLSHSASSLSLQQAASELRRAOMTEGPTAPNPF 1080
 Db 1021 RLRSKVPSEITITVAASTQSPQMLSHSASSLSLQQAASELRRAOMTEGPTAPNPF 1080
 QY 1081 SHGTPPEVVPPLFSSINGVPTAAATAVPATSSPPNDISTSVYQSEVTPTEEGIAGV 1140
 Db 1081 SHGTPPEVVPPLFSSINGVPTAAATAVPATSSPPNDISTSVYQSEVTPTEEGIAGV 1140
 QY 1141 ATSTGVVTSGLPIPPVSESPVLSVSSITIPAVVSISTSPSIQVPTSEIIVSSTA 1200
 Db 1141 ATSTGVVTSGLPIPPVSESPVLSVSSITIPAVVSISTSPSIQVPTSEIIVSSTA 1200
 QY 1201 LYPSTVATSAAGSGATATGPKPAPVAVSQOAGSTTVGATLTVSVSTTTFPSTASDLS 1260
 Db 1201 LYPSTVATSAAGSGATATGPKPAPVAVSQOAGSTTVGATLTVSVSTTTFPSTASDLS 1260
 QY 1261 IQLSSSTSTPTLAETTVVSAHSLDKTSHSITGLAFSLAPSSSSPGAGVSSYISQPGG 1320

Db 1261 IQLSSSTSTPTLAETTVVSAHSLDKTSHSITGLAFSLAPSSSSPGAGVSSYISQPGG 1320
 QY 1321 LHPVLVPSVASTPILPOAGPSTPILPOVPSIPPLVQVAVANVPAVQOGLIHSQPOPAL 1380
 Db 1321 LHPVLVPSVASTPILPOAGPSTPILPOVPSIPPLVQVAVANVPAVQOGLIHSQPOPAL 1380
 QY 1381 LHPQPTHCEVNDTOPKAPGIDDIKTEBKRLSIFSEHSSGAGASVLESTLVIES 1440
 Db 1381 LHPQPTHCEVNDTOPKAPGIDDIKTEBKRLSIFSEHSSGAGASVLESTLVIES 1440
 QY 1441 TTPPGIPTTAVASKILSTSTSTCLPPTNIPLGVALPVTPVVTPOGVSTPVSSTTSGVK 1500
 Db 1441 TTPPGIPTTAVASKILSTSTSTCLPPTNIPLGVALPVTPVVTPOGVSTPVSSTTSGVK 1500
 QY 1501 PGTAPSKPPLTKAPVLVPGVELPAGTLPBQPLPPPPSPSTTOSQOQPLEDDAQLRRLTSP 1560
 Db 1501 PGTAPSKPPLTKAPVLVPGVELPAGTLPBQPLPPPPSPSTTOSQOQPLEDDAQLRRLTSP 1560
 QY 1561 EXITTVSANGPVSMAPATAITEAGTOPKGVSVQKGPVLATSSGAGVFXKGRFQVSA 1620
 Db 1561 EXITTVSANGPVSMAPATAITEAGTOPKGVSVQKGPVLATSSGAGVFXKGRFQVSA 1620
 QY 1621 DGAQKEGKNSDEDAKSVHFESSSTSESVLSSSPESTLVKPEPENGITIPGISTVPESSAH 1680
 Db 1621 DGAQKEGKNSDEDAKSVHFESSSTSESVLSSSPESTLVKPEPENGITIPGISTVPESSAH 1680
 QY 1681 KTTASEAKSDTGOPTKGRFOVTTTANKVGRFVSUKTEDKITDTKKEGPVAPPPFMDLEQ 1740
 Db 1681 KTTASEAKSDTGOPTKGRFOVTTTANKVGRFVSUKTEDKITDTKKEGPVAPPPFMDLEQ 1740
 QY 1741 AVLPVAVIPKKEKELSEPHSHANGPSSDPBEAFLSRVDGSGSPHSQSLSPQNTL 1800
 Db 1741 AVLPVAVIPKKEKELSEPHSHANGPSSDPBEAFLSRVDGSGSPHSQSLSPQNTL 1800
 QY 1801 SOSISNSFNSSNSMSDNESDIEDDELKLEIRLRDLKELKIOLOQSQKHEIBELTYKLG 1860
 Db 1801 SOSISNSFNSSNSMSDNESDIEDDELKLEIRLRDLKELKIOLOQSQKHEIBELTYKLG 1860
 QY 1861 KVPNAVITPPAAPISGRRRPTKSKGSSRSASLGNKSPQLSGNLSGQSAASVLRPOQT 1920
 Db 1861 KVPNAVITPPAAPISGRRRPTKSKGSSRSASLGNKSPQLSGNLSGQSAASVLRPOQT 1920
 QY 1921 LHPGNIPBESQOQLQPLKPSBSSDNLVYAFSTDAISVPSLSAPQOG---TSSTN 1974
 Db 1921 LHPGNIPBESQOQLQPLKPSBSSDNLVYAFSTDAISVPSLSAPQOG---TSSTN 1974

RESULT 14
 US-09-854-856-10
 ; Sequence 10, Application US/09854856
 ; Patent No. 6541252
 ; GENERAL INFORMATION:
 ; APPLICANT: Walke, D. Wade
 ; APPLICANT: Halbun, Erin
 ; APPLICANT: Donoho, Gregory
 ; APPLICANT: Turner, C. Alexander Jr.
 ; TITLE OF INVENTION: No. 6541252el Human Kinases and Polynucleotides
 ; TITLE OF INVENTION: Encoding the Same
 ; FILE REFERENCE: LEX-0178-USA
 ; CURRENT APPLICATION NUMBER: US/09/854,856
 ; PRIOR FILING DATE: 2001-05-14
 ; PRIOR FILING DATE: 2000-05-19
 ; NUMBER OF SEQ ID NOS: 64
 ; SOFTWARE: PasteSeq for Windows Version 4.0
 ; SEQ ID NO 10
 ; LENGTH: 2092
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: VARIANT
 ; LOCATION: (1)...(2092)

OTHER INFORMATION: Xaa = Any Amino Acid
US-09-854-856-10

Query Match 91.3%; Score 9872.5; DB 4; Length 2092;
Best Local Similarity 95.2%; Pred. No. 0;
Matches 1972; Conservative 1; Mismatches 1; Indels 97; Gaps 2;

QY 1 MSGAAEQQSSPTGSLFLSPAPAPKNGSSSDSSVGEELGAAADATVGRTEERRRHT 60
DB 1 MSGAAEQQSSPTGSLFLSPAPAPKNGSSSDSSVGEELGAAADATVGRTEERRRHT 60
QY 61 MDKDSRGAATTTTTEHRRFFRRSVICDSNATALELPGPLSLPQSPISPAAPQSAPEPH 120
DB 61 MDKDSRGAATTTTTEHRRFFRRSVICDSNATALELPGPLSLPQSPISPAAPQSAPEPH 120
QY 121 REEVNTATSOVAAQPPAAAAPGQAVAGAPSTVPSSTKDRPVSOPLSVSKEEPP 180
DB 121 REEVNTATSOVAAQPPAAAAPGQAVAGAPSTVPSSTKDRPVSOPLSVSKEEPP 180
QY 181 ARSGGGGSAKEPQBERSQQODDIELEETKAVGMSNDRFLKPDIEIGRGSFKTVYKGLD 240
DB 181 ARSGGGGSAKEPQBERSQQODDIELEETKAVGMSNDRFLKPDIEIGRGSFKTVYKGLD 240
QY 241 TETTVAVAMCELOQRKLTSEERQRFKEBAEMLKGLQHPNIVRFYDSWESTVKGKCTIVL 300
DB 241 TETTVAVAMCELOQRKLTSEERQRFKEBAEMLKGLQHPNIVRFYDSWESTVKGKCTIVL 300
QY 301 TELMTSGTLKTYLAKRFVYMKVYLRSMCRQLKGLQFHTPTPTIHRDLKCDNIFTGP 360
DB 301 TELMTSGTLKTYLAKRFVYMKVYLRSMCRQLKGLQFHTPTPTIHRDLKCDNIFTGP 360
QY 361 TGSVKIGDLGLATLKRAFAKSVIGTPEFMAPEMYEKEKYDESUVVAFGCMLEMAISEY 420
DB 361 TGSVKIGDLGLATLKRAFAKSVIGTPEFMAPEMYEKEKYDESUVVAFGCMLEMAISEY 420
QY 421 PYSECQNAAOIYRRVTSGVYKPAFADKVAIPVKEIIEGCIKONDERYSIKDLINHAFFQ 480
DB 421 PYSECQNAAOIYRRVTSGVYKPAFADKVAIPVKEIIEGCIKONDERYSIKDLINHAFFQ 480
QY 481 EETGVARELAEBEDGEEKIAIKLMRIDIKLKAKKYKONDAIEFSFPLERDPEDVAQEM 540
DB 481 EETGVARELAEBEDGEEKIAIKLMRIDIKLKAKKYKONDAIEFSFPLERDPEDVAQEM 540
QY 541 VESGVYCEGDHKTAKAIKDRVSLIKRRREOROLVREEOEKKEKESSLKQOVQSSASQ 600
DB 541 VESGVYCEGDHKTAKAIKDRVSLIKRRREOROLVREEOEKKEKESSLKQOVQSSASQ 600
QY 601 TGIKOLPSASTGIPASTTSASVSTQVEPEPEADQHQLOQOQPSISVLSDTGTVDSGQG 660
DB 601 TGIKOLPSASTGIPASTTSASVSTQVEPEPEADQHQLOQOQPSISVLSDTGTVDSGQG 660
QY 661 SSVFTESRVSSQOTVSGSQHEQAHSTGYVGHIPSTVQAQSOQPHGYVPPSSVAQGSQSG 720
DB 661 SSVFTESRVSSQOTVSGSQHEQAHSTGYVGHIPSTVQAQSOQPHGYVPPSSVAQGSQSG 720
QY 721 QPSSSILTVSSSQPIQHPQOQOGIIOOTAPPOOTVOYSLSTOTSSEATTAAPVSPQAP 780
DB 721 QPSSSILTVSSSQPIQHPQOQOGIIOOTAPPOOTVOYSLSTOTSSEATTAAPVSPQAP 780
QY 781 QVLPVUSAGQ-----791
DB 781 QVLPVUSAGQ-----791
QY 792 -----STQGSQVAPAEVAV 807
DB 792 -----STQGSQVAPAEVAV 807
QY 841 VVQPYVESNLVPMGVGVQVQVQPGSGLAQAPITSSQAVLESTQGSQVAPAEVAV 900
DB 841 VVQPYVESNLVPMGVGVQVQVQPGSGLAQAPITSSQAVLESTQGSQVAPAEVAV 900
QY 808 ACPQATQPTTLASSVDSAHSDVAGSMGDNENYPSSSGRHEGRTTKRHYKSVRSRHE 867
DB 808 ACPQATQPTTLASSVDSAHSDVAGSMGDNENYPSSSGRHEGRTTKRHYKSVRSRHE 867
QY 901 ACPQATQPTTLASSVDSAHSDVAGSMGDNENYPSSSGRHEGRTTKRHYKSVRSRHE 960
DB 901 ACPQATQPTTLASSVDSAHSDVAGSMGDNENYPSSSGRHEGRTTKRHYKSVRSRHE 960
QY 868 KTSRPKRLINVSNKGDRAVEQLETHNRKMTVFKEDLDGDNPEELATIMVNNDFILALE 927
DB 868 KTSRPKRLINVSNKGDRAVEQLETHNRKMTVFKEDLDGDNPEELATIMVNNDFILALE 927

DB 961 KTSRPKRLINVSNKGDRAVEQLETHNRKMTVFKEDLDGDNPEELATIMVNNDFILALE 1020
QY 928 RESFVDQVREIIEKADENLSEDSVEPEGQGLSLQKDDYGFSGSKLEGEFFKQPIPA 987
DB 928 RESFVDQVREIIEKADENLSEDSVEPEGQGLSLQKDDYGFSGSKLEGEFFKQPIPA 987
QY 1021 RESFVDQVREIIEKADENLSEDSVEPEGQGLSLQKDDYGFSGSKLEGEFFKQPIPA 1080
DB 1021 RESFVDQVREIIEKADENLSEDSVEPEGQGLSLQKDDYGFSGSKLEGEFFKQPIPA 1080
QY 988 SSMFQOQIGIPSSLTQVYVHSGRRFIVSPVESRLRSKVPFSEITDTVAASTQSPGMN 1047
DB 988 SSMFQOQIGIPSSLTQVYVHSGRRFIVSPVESRLRSKVPFSEITDTVAASTQSPGMN 1047
QY 1048 LSHSASISLQQAASELRRAQMTGPTAPNPNGTGFVVPVPELSSIGVPTTAAT 1107
DB 1048 LSHSASISLQQAASELRRAQMTGPTAPNPNGTGFVVPVPELSSIGVPTTAAT 1107
QY 1141 LSHSASISLQQAASELRRAQMTGPTAPNPNGTGFVVPVPELSSIGVPTTAAT 1200
DB 1141 LSHSASISLQQAASELRRAQMTGPTAPNPNGTGFVVPVPELSSIGVPTTAAT 1200
QY 1108 APVPATSSPNDISTSVIIOSEVTVPEEGIAGVASTGVVTSGGPLPVPVSESPVLSVV 1167
DB 1108 APVPATSSPNDISTSVIIOSEVTVPEEGIAGVASTGVVTSGGPLPVPVSESPVLSVV 1167
QY 1201 APVPATSSPNDISTSVIIOSEVTVPEEGIAGVASTGVVTSGGPLPVPVSESPVLSVV 1260
DB 1201 APVPATSSPNDISTSVIIOSEVTVPEEGIAGVASTGVVTSGGPLPVPVSESPVLSVV 1260
QY 1168 SSITIPAVSISTSPLOVPTSTSEIVVSTALYPSVTSATASAGGSTATGKPRPA 1227
DB 1168 SSITIPAVSISTSPLOVPTSTSEIVVSTALYPSVTSATASAGGSTATGKPRPA 1227
QY 1261 SSITIPAVSISTSPLOVPTSTSEIVVSTALYPSVTSATASAGGSTATGKPRPA 1320
DB 1261 SSITIPAVSISTSPLOVPTSTSEIVVSTALYPSVTSATASAGGSTATGKPRPA 1320
QY 1228 VVSQQAAGSTTVGATLTSVSTTSFPSTASQLSLOLSSSTPTLAETVVVSAHSLDKTS 1287
DB 1228 VVSQQAAGSTTVGATLTSVSTTSFPSTASQLSLOLSSSTPTLAETVVVSAHSLDKTS 1287
QY 1321 VVSQQAAGSTTVGATLTSVSTTSFPSTASQLSLOLSSSTPTLAETVVVSAHSLDKTS 1380
DB 1321 VVSQQAAGSTTVGATLTSVSTTSFPSTASQLSLOLSSSTPTLAETVVVSAHSLDKTS 1380
QY 1288 HSSITGLAFSLAPSSSSSPGAGVSYISQPGIHLPLVPSVIASTPLDQAGPTSTPL 1347
DB 1288 HSSITGLAFSLAPSSSSSPGAGVSYISQPGIHLPLVPSVIASTPLDQAGPTSTPL 1347
QY 1381 HSSITGLAFSLAPSSSSSPGAGVSYISQPGIHLPLVPSVIASTPLDQAGPTSTPL 1440
DB 1381 HSSITGLAFSLAPSSSSSPGAGVSYISQPGIHLPLVPSVIASTPLDQAGPTSTPL 1440
QY 1348 LPQVPSIPPLVQVAVNPAVQOQLIHSQOPALLPQPHHCPEVSDTQPKAGIDIDIX 1407
DB 1348 LPQVPSIPPLVQVAVNPAVQOQLIHSQOPALLPQPHHCPEVSDTQPKAGIDIDIX 1407
QY 1441 LPQVPSIPPLVQVAVNPAVQOQLIHSQOPALLPQPHHCPEVSDTQPKAGIDIDIX 1500
DB 1441 LPQVPSIPPLVQVAVNPAVQOQLIHSQOPALLPQPHHCPEVSDTQPKAGIDIDIX 1500
QY 1408 TLEBKRLSLFSEHSSGQAQASVLETSVLESTVTGIPPTAAVAPSKLITSTSTCLPP 1467
DB 1408 TLEBKRLSLFSEHSSGQAQASVLETSVLESTVTGIPPTAAVAPSKLITSTSTCLPP 1467
QY 1501 TLEBKRLSLFSEHSSGQAQASVLETSVLESTVTGIPPTAAVAPSKLITSTSTCLPP 1560
DB 1501 TLEBKRLSLFSEHSSGQAQASVLETSVLESTVTGIPPTAAVAPSKLITSTSTCLPP 1560
QY 1468 TNLPLGLVALPVPVTPGVSTVSTTSVGVKGPAPSKPLTKADVLPVGTLPAGTL 1527
DB 1468 TNLPLGLVALPVPVTPGVSTVSTTSVGVKGPAPSKPLTKADVLPVGTLPAGTL 1527
QY 1561 TNLPLGLVALPVPVTPGVSTVSTTSVGVKGPAPSKPLTKADVLPVGTLPAGTL 1620
DB 1561 TNLPLGLVALPVPVTPGVSTVSTTSVGVKGPAPSKPLTKADVLPVGTLPAGTL 1620
QY 1528 PSEQLPFPFPGPSLTQSQOPLLEDLDQRLTSLPEKLTIVTSVAVGVSAAPATLATEACTOP 1587
DB 1528 PSEQLPFPFPGPSLTQSQOPLLEDLDQRLTSLPEKLTIVTSVAVGVSAAPATLATEACTOP 1587
QY 1621 PSEQLPFPFPGPSLTQSQOPLLEDLDQRLTSLPEKLTIVTSVAVGVSAAPATLATEACTOP 1680
DB 1621 PSEQLPFPFPGPSLTQSQOPLLEDLDQRLTSLPEKLTIVTSVAVGVSAAPATLATEACTOP 1680
QY 1588 QKGVSVQYKEGPVLTATSGAGVFKMGRFQVVAADGAQKEGKNKSEDAKSVHFEESTSESS 1647
DB 1588 QKGVSVQYKEGPVLTATSGAGVFKMGRFQVVAADGAQKEGKNKSEDAKSVHFEESTSESS 1647
QY 1681 QKGVSVQYKEGPVLTATSGAGVFKMGRFQVVAADGAQKEGKNKSEDAKSVHFEESTSESS 1740
DB 1681 QKGVSVQYKEGPVLTATSGAGVFKMGRFQVVAADGAQKEGKNKSEDAKSVHFEESTSESS 1740
QY 1648 VLSSSPPESTLVKPEPNGITIPGISDVPSAHKTTASEAKSDTGQPTKYGRFOVTTAN 1707
DB 1648 VLSSSPPESTLVKPEPNGITIPGISDVPSAHKTTASEAKSDTGQPTKYGRFOVTTAN 1707
QY 1741 VLSSSPPESTLVKPEPNGITIPGISDVPSAHKTTASEAKSDTGQPTKYGRFOVTTAN 1800
DB 1741 VLSSSPPESTLVKPEPNGITIPGISDVPSAHKTTASEAKSDTGQPTKYGRFOVTTAN 1800
QY 1708 KYGRFVSYSKTEEDITITTKKGPVAPSPFMDLEAVIPAVIPKKEKPELSPSHLNGSSD 1767
DB 1708 KYGRFVSYSKTEEDITITTKKGPVAPSPFMDLEAVIPAVIPKKEKPELSPSHLNGSSD 1767
QY 1801 KYGRFVSYSKTEEDITITTKKGPVAPSPFMDLEAVIPAVIPKKEKPELSPSHLNGSSD 1860
DB 1801 KYGRFVSYSKTEEDITITTKKGPVAPSPFMDLEAVIPAVIPKKEKPELSPSHLNGSSD 1860
QY 1768 PEAAFLSRVDVDSGSPHSPHQLSKSLPSQNLSSQSLSNSFNSYSMSDNESEDIEDDLK 1827
DB 1768 PEAAFLSRVDVDSGSPHSPHQLSKSLPSQNLSSQSLSNSFNSYSMSDNESEDIEDDLK 1827
QY 1861 PEAAFLSRVDVDSGSPHSPHQLSKSLPSQNLSSQSLSNSFNSYSMSDNESEDIEDDLK 1920
DB 1861 PEAAFLSRVDVDSGSPHSPHQLSKSLPSQNLSSQSLSNSFNSYSMSDNESEDIEDDLK 1920
QY 1921 LEIRRLRDHKLKAIQDLOSQKHEIESLYTKLGKVPAPVAVIPPAALSGRRRPTKSGS 1980
DB 1921 LEIRRLRDHKLKAIQDLOSQKHEIESLYTKLGKVPAPVAVIPPAALSGRRRPTKSGS 1980
QY 1888 KSSRSSSLGKNSQULSGNLSGQSAASVULHPOQLTHPQNTPESGOQLOLQPLKSPSSDN 1947
DB 1888 KSSRSSSLGKNSQULSGNLSGQSAASVULHPOQLTHPQNTPESGOQLOLQPLKSPSSDN 1947
QY 1981 KSSRSSSLGKNSQULSGNLSGQSAASVULHPOQLTHPQNTPESGOQLOLQPLKSPSSDN 2040
DB 1981 KSSRSSSLGKNSQULSGNLSGQSAASVULHPOQLTHPQNTPESGOQLOLQPLKSPSSDN 2040
QY 1948 LYSAPTSQDGLSVPSLSAPGQG-----TSSSTN 1974
DB 1948 LYSAPTSQDGLSVPSLSAPGQG-----TSSSTN 1974
QY 2041 LYSAPTSQDGLSVPSLSAPGQG-----TSSSTN 2071
DB 2041 LYSAPTSQDGLSVPSLSAPGQG-----TSSSTN 2071

RESULT 15
US-09-854-856-4
Sequence 4, Application US/09854856
Patent No. 6541252
GENERAL INFORMATION:
APPLICANT: Walke, D. Wade
APPLICANT: Hilbun, Erin
APPLICANT: Donoho, Gregory
APPLICANT: Turner, C. Alexander Jr.
TITLE OF INVENTION: No. 6541252el Human Kinases and Polynucleotides
FILE REFERENCE: LEX-0178-USA
CURRENT APPLICATION NUMBER: US/09/854,856
CURRENT FILING DATE: 2001-05-14
PRIORITY FILING DATE: 2000-05-19
PRIORITY FILING DATE: 2000-05-19
NUMBER OF SEQ ID NOS: 64
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 4
LENGTH: 2245
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: VARIANT
LOCATION: (1)-(2245)
OTHER INFORMATION: Xaa = Any Amino Acid
US-09-854-856-4

Query Match 90.6%; Score 9796; DB 4; Length 2245;
Best Local Similarity 88.7%; Pred. No. 0;
Matches 1972; Conservative 1; Mismatches 1; Indels 250; Gaps 2;

QY 1 MSGGAAGKSSSTPGSLFLSPAPAPKKGSSSDSVGKLGAAADAYTGTEYRRRRHT 60
DB 1 MSGGAAGKSSSTPGSLFLSPAPAPKKGSSSDSVGKLGAAADAYTGTEYRRRRHT 60
QY 61 MDKSRBAANTTTTHERFPFRSVICDSNATLELPG.LPLSPPOSIPAAVPOGAPPEPH 120
DB 61 MDKSRBAANTTTTHERFPFRSVICDSNATLELPG.LPLSPPOSIPAAVPOGAPPEPH 120
QY 121 REETVATATSQVAQPPAAAPGQAVAGPAPSTVSTSKDRAVPSOPSLVSGKEBPP 180
DB 121 REETVATATSQVAQPPAAAPGQAVAGPAPSTVSTSKDRAVPSOPSLVSGKEBPP 180
QY 181 ARSGSGGSAKEPQERSQODDIELEETKAVGMSNDGRFLKFDIEIGRGSFKTVYKGLD 240
DB 181 ARSGSGGSAKEPQERSQODDIELEETKAVGMSNDGRFLKFDIEIGRGSFKTVYKGLD 240
QY 241 TETTVAVAMCLOBRKLTLSRRORFKERAEMLKGLPHNIVRFYDSMESTVKGKCIYLV 300
DB 241 TETTVAVAMCLOBRKLTLSRRORFKERAEMLKGLPHNIVRFYDSMESTVKGKCIYLV 300
QY 301 TELMTSGTLTKYIKRFKMKIKVLRWCROILKGLQFLHTRTPLIHRDLKCDNIFITGP 360
DB 301 TELMTSGTLTKYIKRFKMKIKVLRWCROILKGLQFLHTRTPLIHRDLKCDNIFITGP 360
QY 361 TGSVKTIGDLATLKRAFPKSVIGTPEFMAPEMYEEKYDESUVYVAGMCMLEMAISEY 420
DB 361 TGSVKTIGDLATLKRAFPKSVIGTPEFMAPEMYEEKYDESUVYVAGMCMLEMAISEY 420
QY 421 PYSCQNAAOIYRVYTSVGVKSPKVAIPVKEITIEGCIKONDERYSIDOLNHAFFQ 480
DB 421 PYSCQNAAOIYRVYTSVGVKSPKVAIPVKEITIEGCIKONDERYSIDOLNHAFFQ 480
QY 481 EETGVRAVLAEDDEGEKIAIKLWLRITEDIKLGKGYKONDEAIERSFDLERDVPEDVAGEM 540
DB 481 EETGVRAVLAEDDEGEKIAIKLWLRITEDIKLGKGYKONDEAIERSFDLERDVPEDVAGEM 540
QY 541 VESGVGCEGDKHTMAKAIKRVSLIKRGRGQOLVREBQEKKKQVESSLKQOVESGASQ 600
DB 541 VESGVGCEGDKHTMAKAIKRVSLIKRGRGQOLVREBQEKKKQVESSLKQOVESGASQ 600

QY 601 TGIKOLPSASTGIPITASTTSASVSTQVEBEPBADOHQLOIQOQPSISVLSDGTVNSGQ 660
DB 601 TGIKOLPSASTGIPITASTTSASVSTQVEBEPBADOHQLOIQOQPSISVLSDGTVNSGQ 660
QY 661 SSYFTSRVSSQCTVYSGQHEQAHSTGVPHGIPSTVQAQSPHGVYPPSSVAQSQSQG 720
DB 661 SSYFTSRVSSQCTVYSGQHEQAHSTGVPHGIPSTVQAQSPHGVYPPSSVAQSQSQG 720
QY 721 QPSSSLTGVSNSOPIQHPQOQOQGIQOTAPPOQTVQISQTSSEATTAQVPSQPAQ 780
DB 721 QPSSSLTGVSNSOPIQHPQOQOQGIQOTAPPOQTVQISQTSSEATTAQVPSQPAQ 780
QY 781 QVLPQVSAGKO----- 791
DB 781 QVLPQVSAGKO----- 791
QY 792 ----- 791
DB 792 ----- 791
QY 841 VSGIPSTPHVSTAQTGFSSLPITMAAGITOPLLTASATTAIRGVSTVPSQLPTLL 900
DB 841 VSGIPSTPHVSTAQTGFSSLPITMAAGITOPLLTASATTAIRGVSTVPSQLPTLL 900
QY 901 QPVTQLPSQVHPOLLQPAVQSMGIPANLGOAAEVLPSGQVLYXGPPRRLPQYPGDSNI 960
DB 901 QPVTQLPSQVHPOLLQPAVQSMGIPANLGOAAEVLPSGQVLYXGPPRRLPQYPGDSNI 960
QY 992 ----- 791
DB 992 ----- 791
QY 1021 GSIAQAPTTSSQQAVALSTQGVQVAPAEVAAVAAQPAQPTPTLLASSVDASHSVASGMS 1080
DB 1021 GSIAQAPTTSSQQAVALSTQGVQVAPAEVAAVAAQPAQPTPTLLASSVDASHSVASGMS 1080
QY 835 DGNENVPSSGRHGRRTTKRKYRSVRSRSHKTSRPKRIINVSNGRVRVECOLETH 894
DB 835 DGNENVPSSGRHGRRTTKRKYRSVRSRSHKTSRPKRIINVSNGRVRVECOLETH 894
QY 1081 DGNENVPSSGRHGRRTTKRKYRSVRSRSHKTSRPKRIINVSNGRVRVECOLETH 1140
DB 1081 DGNENVPSSGRHGRRTTKRKYRSVRSRSHKTSRPKRIINVSNGRVRVECOLETH 1140
QY 895 NRKQVTFKFDLDGNDPEIATIMWNDFIILIERESFVQVRETIIEKADMLSEDEVPE 954
DB 895 NRKQVTFKFDLDGNDPEIATIMWNDFIILIERESFVQVRETIIEKADMLSEDEVPE 954
QY 1141 NRKQVTFKFDLDGNDPEIATIMWNDFIILIERESFVQVRETIIEKADMLSEDEVPE 1200
DB 1141 NRKQVTFKFDLDGNDPEIATIMWNDFIILIERESFVQVRETIIEKADMLSEDEVPE 1200
QY 955 EGDQGLSLQKQDYGSGSQKIEGFKQIIPASSYQOIGIPSSITQVHSGRRFTV 1014
DB 955 EGDQGLSLQKQDYGSGSQKIEGFKQIIPASSYQOIGIPSSITQVHSGRRFTV 1014
QY 1201 EGDQGLSLQKQDYGSGSQKIEGFKQIIPASSYQOIGIPSSITQVHSGRRFTV 1260
DB 1201 EGDQGLSLQKQDYGSGSQKIEGFKQIIPASSYQOIGIPSSITQVHSGRRFTV 1260
QY 1015 SPVSESRRESKVPSEITDTVAASQAQSGMMLSHSASSLSLQAAQSELRRAQMTGEPN 1074
DB 1015 SPVSESRRESKVPSEITDTVAASQAQSGMMLSHSASSLSLQAAQSELRRAQMTGEPN 1074
QY 1261 SPVSESRRESKVPSEITDTVAASQAQSGMMLSHSASSLSLQAAQSELRRAQMTGEPN 1320
DB 1261 SPVSESRRESKVPSEITDTVAASQAQSGMMLSHSASSLSLQAAQSELRRAQMTGEPN 1320
QY 1075 TAPNFSHTGPTPEVVPFPLSSIAGVPTTAATAIPVPAISSPPNDISTSVIQSEVTVPT 1134
DB 1075 TAPNFSHTGPTPEVVPFPLSSIAGVPTTAATAIPVPAISSPPNDISTSVIQSEVTVPT 1134
QY 1135 EGIAGVATSTGVNTSGGLPIIPVSESPVLSVSVSITIPVAVSITSTSPSLQVFTSSEI 1194
DB 1135 EGIAGVATSTGVNTSGGLPIIPVSESPVLSVSVSITIPVAVSITSTSPSLQVFTSSEI 1194
QY 1381 EGIAGVATSTGVNTSGGLPIIPVSESPVLSVSVSITIPVAVSITSTSPSLQVFTSSEI 1440
DB 1381 EGIAGVATSTGVNTSGGLPIIPVSESPVLSVSVSITIPVAVSITSTSPSLQVFTSSEI 1440
QY 1195 VVSSTALYPSVTSVATSSAGSGTATGPKPRAVVSQQAAGSTTVGATLTVSFTTSPS 1254
DB 1195 VVSSTALYPSVTSVATSSAGSGTATGPKPRAVVSQQAAGSTTVGATLTVSFTTSPS 1254
QY 1441 VVSSTALYPSVTSVATSSAGSGTATGPKPRAVVSQQAAGSTTVGATLTVSFTTSPS 1500
DB 1441 VVSSTALYPSVTSVATSSAGSGTATGPKPRAVVSQQAAGSTTVGATLTVSFTTSPS 1500
QY 1315 ISQGLHPLVIBSVIASTPILPQAAGPTSPILLPQVPSIPPLVQPVANVAVOQTLIHS 1374
DB 1315 ISQGLHPLVIBSVIASTPILPQAAGPTSPILLPQVPSIPPLVQPVANVAVOQTLIHS 1374
QY 1375 ISQGLHPLVIBSVIASTPILPQAAGPTSPILLPQVPSIPPLVQPVANVAVOQTLIHS 1434
DB 1375 ISQGLHPLVIBSVIASTPILPQAAGPTSPILLPQVPSIPPLVQPVANVAVOQTLIHS 1434
QY 1621 QPQALPBNQHTHCPEVSDTOPKAPGIDIDIKLEKLRSLFSEHSSSGAQAHSVSLT 1680
DB 1621 QPQALPBNQHTHCPEVSDTOPKAPGIDIDIKLEKLRSLFSEHSSSGAQAHSVSLT 1680

QY 1435 SLVIESTVTPGIPPTTAVAPSKLITSTSTCLPPTNLPLGTVALPVTVPVTPGOVSTPVST 1494
 Db 1681 SLVIESTVTPGIPPTTAVAPSKLITSTSTCLPPTNLPLGTVALPVTVPVTPGOVSTPVST 1740
 QY 1495 TTSGVYKPGTAPSKPPLITAPVLPVGTBLPAGTLPEEQLPPPPGSLTOSQOPLIEDLDAQL 1554
 Db 1741 TTSGVYKPGTAPSKPPLITAPVLPVGTBLPAGTLPEEQLPPPPGSLTOSQOPLIEDLDAQL 1800
 QY 1555 RRTLSPEKITTTSAGVPVMAAPTAITEAGTQPOKGVQVQVEGVPYATSSGAGVFKMGF 1614
 Db 1801 RRTLSPEKITTTSAGVPVMAAPTAITEAGTQPOKGVQVQVEGVPYATSSGAGVFKMGF 1860
 QY 1615 QVSVAADGAQOKEGKNKSEDAKSVHESSTSSSVLSSSSPESTLVKPPENGITTPGISSD 1674
 Db 1861 QVSVAADGAQOKEGKNKSEDAKSVHESSTSSSVLSSSSPESTLVKPPENGITTPGISSD 1920
 QY 1675 VPESAHKTTSSEAKSDTQOPTKVGRFQVTTTANKVGRFVSCTEDKITTDTKKEGVPVSP 1734
 Db 1921 VPESAHKTTSSEAKSDTQOPTKVGRFQVTTTANKVGRFVSCTEDKITTDTKKEGVPVSP 1980
 QY 1735 FMDLEQAVLPVAVIPKKEPELSEPSHLNGPSSDPEAAFLSRDNDGSGSPHSPHOLSKS 1794
 Db 1981 FMDLEQAVLPVAVIPKKEPELSEPSHLNGPSSDPEAAFLSRDNDGSGSPHSPHOLSKS 2040
 QY 1795 LPSQNLQSLSNSFNSSSYMSSDNESDIEDDLKLELRRLRDKHKEIQDLQSRQKHIES 1854
 Db 2041 LPSQNLQSLSNSFNSSSYMSSDNESDIEDDLKLELRRLRDKHKEIQDLQSRQKHIES 2100
 QY 1855 LYTKLGKVPVPPVILPPAAPLSGRRRPTKSGSKSSRSSSLGNKSPOLSGNLSCQSAASV 1914
 Db 2101 LYTKLGKVPVPPVILPPAAPLSGRRRPTKSGSKSSRSSSLGNKSPOLSGNLSCQSAASV 2160
 QY 1915 LHPQOQLHPPGNIPESGONQLQPLKPSPSGDNLYSAFTSDGATSVPSLSAPQGG---T 1970
 Db 2161 LHPQOQLHPPGNIPESGONQLQPLKPSPSGDNLYSAFTSDGATSVPSLSAPQGG---T 2220
 QY 1971 SSTN 1974
 Db 2221 SSKN 2224

Search completed: September 24, 2004, 01:12:44
 Job time : 64 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 24, 2004, 00:51:58 ; Search time 189 Seconds

(without alignments)
3193.235 Million cell updates/sec

Title: US-10-010-720-14

Perfect score: 10812

Sequence: 1 MSGGAAXKQSTPGSLFLSP.....NISMLKRSINSPGSLNRTT 2136

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: A_Geneseq_29Jan04:*
2: geneseqp1980s:*
3: geneseqp1990s:*
4: geneseqp2000s:*
5: geneseqp2001s:*
6: geneseqp2002s:*
7: geneseqp2003as:*
8: geneseqp2003bs:*
9: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	10810	100.0	2136	7	ABO44387 Novel hum
2	10794.5	99.8	2135	5	AAE21714 Human PKI
3	10753.5	99.5	2229	7	ABO44384 Novel hum
4	10677	98.8	2382	6	ABP71619 Human WNK
5	10677	98.8	2382	7	ABO44381 Novel hum
6	10646	98.5	2108	7	ABO44395 Novel hum
7	10646	98.5	2108	7	ADCG9049 Novel hum
8	10589.5	97.9	2201	7	ABO44392 Novel hum
9	10513	97.2	2354	7	ABO44389 Novel hum
10	10508	97.2	2076	7	ABO44403 Novel hum
11	10451.5	96.7	2169	7	ABO44400 Novel hum
12	10375	96.0	2332	7	ABO44397 Novel hum
13	10344	95.7	2048	7	ABO44411 Novel hum
14	10287.5	95.1	2141	7	ABO44408 Novel hum
15	10211	94.4	2294	7	ABO44405 Novel hum
16	9929	91.8	1939	7	ABO44388 Novel hum
17	9872.5	91.3	2092	7	ABO44385 Novel hum
18	9796	90.6	2245	7	ABO44382 Novel hum
19	9765	90.3	1971	7	ABO44381 Novel hum
20	9708.5	89.8	2064	7	ABO44396 Novel hum
21	9632	89.1	2217	7	ABO44393 Novel hum
22	9627	89.0	1939	7	ABO44404 Novel hum
23	9612.5	88.9	1920	4	AAE25097 Human kin
24	9570.5	88.5	2032	7	ABO44401 Novel hum
25	9494	87.8	2185	7	ABO44398 Novel hum

26	9463	87.5	1911	7	ABO44412 Novel hum
27	9406.5	87.0	2004	7	ABO44409 Novel hum
28	9330	86.3	2157	7	ABO44406 Novel hum
29	4053	37.5	982	7	ABO44383 Novel hum
30	4043	37.4	829	7	ABO44386 Novel hum
31	3889	36.0	954	7	ABO44391 Novel hum
32	3879	35.9	801	7	ABO44394 Novel hum
33	3751	34.7	922	7	ABO44399 Novel hum
34	3741	34.6	769	7	ABO44402 Novel hum
35	3587	33.2	894	7	ABO44407 Novel hum
36	3577	33.1	741	7	ABO44410 Novel hum
37	2903	26.8	565	4	AAE25097 Human kin
38	2631	24.3	2193	6	ABR4219 Human kin
39	2524	23.3	502	4	AAE25097 Human kin
40	2469.5	22.8	591	4	AAE25097 Human kin
41	2419	22.4	470	4	AAE25097 Human kin
42	2419	22.4	470	4	AAE25097 Human kin
43	2130.5	19.7	1345	5	AAE25097 Human kin
44	2126	19.7	1243	6	ABP71620 Human WNK
45	2126	19.7	1513	4	AAU03531 Human pro

ALIGNMENTS

RESULT 1	
ID	ABO44387 standard; protein; 2136 AA.
AC	ABO44387;
DT	26-SEP-2003 (first entry)
DE	Novel human protein kinase #7.
KW	Human, kinase; enzyme; cosmetic application; nutraceutical application.
OS	Homo sapiens.
PN	US6541252-B1.
PD	01-APR-2003.
PF	14-MAY-2001; 2001US-00854856.
PR	19-MAY-2000; 2000US-0206015P.
PA	(LEXI-) LEXICON GENETICS INC.
PI	Walke DW, Hilbun E, Donoho G, Turner CA;
PI	WPI; 2003-575927/54.
DR	N-PSDB; ACH03770.
XX	
XX	New nucleic acid encoding novel human proteins, useful in cosmetic and nutraceutical applications.
XX	Claim 1; Page; 11pp; English.
XX	The invention relates to a new isolated nucleic acid encoding a novel human protein kinase. The nucleic acid is useful in cosmetic and nutraceutical applications. The present sequence represents the amino acid sequence of a novel human protein kinase. Note: The sequence data for this patent did not from part of the printed specification but was obtained in electronic format directly from USPTO at seqdata.uspto.gov/sequence.html?DocID=6541252B1
SO	Sequence 2136 AA;
Query Match	100.0%; Score 10810; DB 7; Length 2136;
Best Local Similarity	100.0%; Pred. No. 0;
Matches 2136; Conservative	0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSGGAABKOSSTGSLFLSPAPAPKNGSSSDSSVGEKLGAAADAATVGTTEERYRRRHT 60
 DB 1 MSGGAABKOSSTGSLFLSPAPAPKNGSSSDSSVGEKLGAAADAATVGTTEERYRRRHT 60
 QY 61 MDOSKGAATTTTTERFRFRSVI CDSNATALEPLPLPLPOPSPIPAPOSAPPEPH 120
 DB 61 MDOSKGAATTTTTERFRFRSVI CDSNATALEPLPLPLPOPSPIPAPOSAPPEPH 120
 QY 121 REETVATATSVQAQOPPAAPAEQAVAGPADSVTSSTSKORPVSOPELVSKKEPP 180
 DB 121 REETVATATSVQAQOPPAAPAEQAVAGPADSVTSSTSKORPVSOPELVSKKEPP 180
 QY 181 ARSGSGGSKKEPOEERSQOQDDLELETKAVGMSNGRFLKFDIEIGRSFXTVYKGLD 240
 DB 181 ARSGSGGSKKEPOEERSQOQDDLELETKAVGMSNGRFLKFDIEIGRSFXTVYKGLD 240
 QY 241 TETTVAVAMELODRKLTKEERQEFKEAEMLKGLQHPNIVRFYDSESTVKGKCIVLV 300
 DB 241 TETTVAVAMELODRKLTKEERQEFKEAEMLKGLQHPNIVRFYDSESTVKGKCIVLV 300
 QY 301 TELMTSGTLKTYLKRFRVWKIKVLRSWCROILKGLQFLHTRTPPIIHRDLKCNIFITGP 360
 DB 301 TELMTSGTLKTYLKRFRVWKIKVLRSWCROILKGLQFLHTRTPPIIHRDLKCNIFITGP 360
 QY 361 TGSVKIGDLGATLKRAFPKSVITGPEFMAPEMYEEKYDESVDVYAFGKCMLEMATSEY 420
 DB 361 TGSVKIGDLGATLKRAFPKSVITGPEFMAPEMYEEKYDESVDVYAFGKCMLEMATSEY 420
 QY 421 PYSCQNAAOIYRRVTSVGVPAFDPKVAIPEVKEIIEGCIRONKDERYSIKDLNHAFFQ 480
 DB 421 PYSCQNAAOIYRRVTSVGVPAFDPKVAIPEVKEIIEGCIRONKDERYSIKDLNHAFFQ 480
 QY 481 BETGVAVELAEEDDEKCIATKLMRIDIKLKKYKONENAIERSFLBERDVPEDVAQEM 540
 DB 481 BETGVAVELAEEDDEKCIATKLMRIDIKLKKYKONENAIERSFLBERDVPEDVAQEM 540
 QY 541 VESGYVEGKHMAKAIKORVSLIKRREOROLVREOEKKOEKKESSLKQOVQSSASQ 600
 DB 541 VESGYVEGKHMAKAIKORVSLIKRREOROLVREOEKKOEKKESSLKQOVQSSASQ 600
 QY 601 TGIKOLPSASTGIPJTASTSASVTOVEPEEPADQOOLQYQOPIISVLDGTVDGSGG 660
 DB 601 TGIKOLPSASTGIPJTASTSASVTOVEPEEPADQOOLQYQOPIISVLDGTVDGSGG 660
 QY 661 SSVFTESRVSQOQVYSGSOHEQASHSTGVGHIPSTVOAQSOQPHGVYPPSSVAQSOQ 720
 DB 661 SSVFTESRVSQOQVYSGSOHEQASHSTGVGHIPSTVOAQSOQPHGVYPPSSVAQSOQ 720
 QY 721 QPSSSSLTGVSSQPIQHPOQOQSIQOTAPPOQTVQVYSLSTQSTSSSEKTTAQPVSOQAP 780
 DB 721 QPSSSSLTGVSSQPIQHPOQOQSIQOTAPPOQTVQVYSLSTQSTSSSEKTTAQPVSOQAP 780
 QY 781 QVLPOVSAKQOSTQGVSAQVAPAEPAVAVAQPOATQPTLLASSVDASHSDVAGSGDGENV 840
 DB 781 QVLPOVSAKQOSTQGVSAQVAPAEPAVAVAQPOATQPTLLASSVDASHSDVAGSGDGENV 840
 QY 841 PSSSGRHRGRTTKKHRYKSVASRSRHEKTSRPXRIILVNSKNGRVYECQLETNRKVT 900
 DB 841 PSSSGRHRGRTTKKHRYKSVASRSRHEKTSRPXRIILVNSKNGRVYECQLETNRKVT 900
 QY 901 FKPLDGNPEEITIMVNNDFILAIERESFVDQREIIEKADEMLSEDVVEEGDQGL 960
 DB 901 FKPLDGNPEEITIMVNNDFILAIERESFVDQREIIEKADEMLSEDVVEEGDQGL 960
 QY 961 ESLOGKDYFSGSOKLEGEFKOPIPASSMPQOIGIPTSSLTOVHSAGRFIVSPVES 1020
 DB 961 ESLOGKDYFSGSOKLEGEFKOPIPASSMPQOIGIPTSSLTOVHSAGRFIVSPVES 1020
 QY 1021 RLRSKVPFSSITTTVAASTAQSPGMNLSHASSLSLQOAFSELRRAOMTGPNTAPNF 1080
 DB 1021 RLRSKVPFSSITTTVAASTAQSPGMNLSHASSLSLQOAFSELRRAOMTGPNTAPNF 1080
 QY 1081 SHGPTFPVVPPLSSINAGVPTTAAATAPVPATSSFPNDISTSVIQSEVTVTEEGIAGV 1140

DB 1081 SHGPTFPVVPPLSSINAGVPTTAAATAPVPATSSFPNDISTSVIQSEVTVTEEGIAGV 1140
 QY 1141 ANSTGVVTSQGLPIPIPVSESPLYSSVSSITTPAVVISITSSPSLOVPTSTSEIIVSSSTA 1200
 DB 1141 ANSTGVVTSQGLPIPIPVSESPLYSSVSSITTPAVVISITSSPSLOVPTSTSEIIVSSSTA 1200
 QY 1201 LVPSTVVSATASAGGSTATPGKPPAVVSOQAAGSTVWATLTSTVSTTSPSTASQLS 1260
 DB 1201 LVPSTVVSATASAGGSTATPGKPPAVVSOQAAGSTVWATLTSTVSTTSPSTASQLS 1260
 QY 1261 IOLSSSTPTTAEIVVWVSAHSDKTSHSSTGLAFSLASAPSSSSSPGACVSSYISQPG 1320
 DB 1261 IOLSSSTPTTAEIVVWVSAHSDKTSHSSTGLAFSLASAPSSSSSPGACVSSYISQPG 1320
 QY 1321 LHPVLVPSVASTAPPIIPOAAGPSTELLPOVBEIIPPLYOVAVNAVPOQTLHISQOPAL 1380
 DB 1321 LHPVLVPSVASTAPPIIPOAAGPSTELLPOVBEIIPPLYOVAVNAVPOQTLHISQOPAL 1380
 QY 1381 LHPQPHHCPEVSDTOPKAPGIDDKITLBEKLSLSEHSSSGAQAHSVSLSTLVIES 1440
 DB 1381 LHPQPHHCPEVSDTOPKAPGIDDKITLBEKLSLSEHSSSGAQAHSVSLSTLVIES 1440
 QY 1441 TVTPGIPTTAVAPSKLITSTTCLPPTNLPIGTVALPVTVPQVSTPVSTTSGVK 1500
 DB 1441 TVTPGIPTTAVAPSKLITSTTCLPPTNLPIGTVALPVTVPQVSTPVSTTSGVK 1500
 QY 1501 PGTAPEKPELTAKPVLPVGHLEPAGTLPSPOLPPRPBPSITQSOQPLEDDAOLRRFLSP 1560
 DB 1501 PGTAPEKPELTAKPVLPVGHLEPAGTLPSPOLPPRPBPSITQSOQPLEDDAOLRRFLSP 1560
 QY 1561 EXITVTSVAVPVSMAAPATAITEAGTOPQKGVSOVKEGPVLATSSGAVFPMGRFOVSVA 1620
 DB 1561 EXITVTSVAVPVSMAAPATAITEAGTOPQKGVSOVKEGPVLATSSGAVFPMGRFOVSVA 1620
 QY 1621 DGAQKEGKNSEDAKSVHFEESSTSESSVSSSPBSTLVKPEPNGITIPGISDVPEBSAH 1680
 DB 1621 DGAQKEGKNSEDAKSVHFEESSTSESSVSSSPBSTLVKPEPNGITIPGISDVPEBSAH 1680
 QY 1681 KTTASEKASTQOPTKGRFOVTTANKVRFVSUKTEDKITTPKKGPAVSPFMDLBQ 1740
 DB 1681 KTTASEKASTQOPTKGRFOVTTANKVRFVSUKTEDKITTPKKGPAVSPFMDLBQ 1740
 QY 1741 AVLPVAVPKKEKPELSPSHLNGSSDPEAFLSRVDVDSGSPHSFHOLESKSLPSONL 1800
 DB 1741 AVLPVAVPKKEKPELSPSHLNGSSDPEAFLSRVDVDSGSPHSFHOLESKSLPSONL 1800
 QY 1801 SOSLSNSFNSSWYSSDNESEDIEDEDLKLELRRLADKHLKEIODIQRKHEIBSLYTKLG 1860
 DB 1801 SOSLSNSFNSSWYSSDNESEDIEDEDLKLELRRLADKHLKEIODIQRKHEIBSLYTKLG 1860
 QY 1861 KVPRAVILPPAALPSGRRRTYKSGKSRSRSGSLGKNSPOLSGNLSGQSAASYLHPQOT 1920
 DB 1861 KVPRAVILPPAALPSGRRRTYKSGKSRSRSGSLGKNSPOLSGNLSGQSAASYLHPQOT 1920
 QY 1921 LHPGNTPESSQONLOPLKPSPESSDNLVSAFTSDGALSVPSLSAPQGTSTNTVGAIV 1980
 DB 1921 LHPGNTPESSQONLOPLKPSPESSDNLVSAFTSDGALSVPSLSAPQGTSTNTVGAIV 1980
 QY 1981 NSQAQAQOPAMTSSRRGTFTDDHLKLVDMVARDAMNLSGRGSGKHMNYGPGMARKFS 2040
 DB 1981 NSQAQAQOPAMTSSRRGTFTDDHLKLVDMVARDAMNLSGRGSGKHMNYGPGMARKFS 2040
 QY 2041 APGOLCISMTSNLGGSAPIASAATSILGHFTKSMCPQOIGFPAITPFGAOMSGTGAPAPQ 2100
 DB 2041 APGOLCISMTSNLGGSAPIASAATSILGHFTKSMCPQOIGFPAITPFGAOMSGTGAPAPQ 2100
 QY 2101 PLGQFOPVGTASLONFINISNLOKSIINPPGSGNLRTT 2136
 DB 2101 PLGQFOPVGTASLONFINISNLOKSIINPPGSGNLRTT 2136

RESULT 2


```

Db      840 PSSGRHGRITTKHRYKSVSRSHETKSRPKRLILNVSNKGRVVECOLETHNRKMYT 899
Qy      901 FKFDLDGDNPEELIATIMVNDNFILAIERESFVDQVREIIIEKADEMLSEDSVEEGDGL 960
Db      900 FKFDLDGDNPEELIATIMVNDNFILAIERESFVDQVREIIIEKADEMLSEDSVEEGDGL 959
Qy      961 ESLOGKDYGSGSGCKEGERKOPIPASMPQOIGIPSSSLTQVHSGRFFIVSPVES 1020
Db      960 ESLOGKDYGSGSGCKEGERKOPIPASMPQOIGIPSSSLTQVHSGRFFIVSPVES 1019
Qy      1021 RLRSKVPSEIITDTVAASTAQSPGMNLSHSASISLQOAFSELRRAOMTEGPTAPNPF 1080
Db      1020 RLRSKVPSEIITDTVAASTAQSPGMNLSHSASISLQOAFSELRRAOMTEGPTAPNPF 1079
Qy      1081 SHTGPTFPVVPFPLSSLAGVTTAAATAPVATSPENDISTSVIQSEVTVPTBEGLAGV 1140
Db      1080 SHTGPTFPVVPFPLSSLAGVTTAAATAPVATSPENDISTSVIQSEVTVPTBEGLAGV 1139
Qy      1141 ATSTGVTSAGGLPIPVSESPVLSVSGSITIPAVSISTSPSQVPTSTSEIIVSSTA 1200
Db      1140 ATSTGVTSAGGLPIPVSESPVLSVSGSITIPAVSISTSPSQVPTSTSEIIVSSTA 1139
Qy      1201 LYPSTVTSATSAAGSITATGPKPPAVVSOQAAGSTTGAATLTSVSTTSFPSTASQLS 1260
Db      1200 LYPSTVTSATSAAGSITATGPKPPAVVSOQAAGSTTGAATLTSVSTTSFPSTASQLS 1259
Qy      1261 IQLSSTSTPTLAETVTVVSAHSLDKTSHSSTTGCLAFSLASPSSSSPGAGVSIYSGPGG 1320
Db      1260 IQLSSTSTPTLAETVTVVSAHSLDKTSHSSTTGCLAFSLASPSSSSPGAGVSIYSGPGG 1319
Qy      1321 LHPVIVPVIASPTPLIPQAAGPTSTPLIPQVSIPIVQVPAVPAVQOQLIHSGOPAL 1380
Db      1320 LHPVIVPVIASPTPLIPQAAGPTSTPLIPQVSIPIVQVPAVPAVQOQLIHSGOPAL 1379
Qy      1381 LPNQPTHCPEDSDTOPKAFGIDIKITLSEKLSLFEHSSSGAQHSAVSLSTLVIES 1440
Db      1380 LPNQPTHCPEDSDTOPKAFGIDIKITLSEKLSLFEHSSSGAQHSAVSLSTLVIES 1439
Qy      1441 TVTNGIPTTANAPSKLITSTTCLPPTNMLPLGVALPVTVVTVPGQVSTVSTTSGVK 1500
Db      1440 TVTNGIPTTANAPSKLITSTTCLPPTNMLPLGVALPVTVVTVPGQVSTVSTTSGVK 1499
Qy      1501 PGTAPEKPLTKAPLPIVTELPAGTLPSEQLPPPPGSLTQSQOPELDDLAOLRRTLSP 1560
Db      1500 PGTAPEKPLTKAPLPIVTELPAGTLPSEQLPPPPGSLTQSQOPELDDLAOLRRTLSP 1559
Qy      1561 EXITVTSAVGVSAAPATITEAGTOPQKGVSOVKEGVLATSSGAGVFEMKGRFQVSVA 1620
Db      1560 EXITVTSAVGVSAAPATITEAGTOPQKGVSOVKEGVLATSSGAGVFEMKGRFQVSVA 1619
Qy      1621 DGAQKEGKNKEDAKSVHFESESTSESSVLSSSSPBSTLYKEEPNGITTPGISSDVPESAH 1680
Db      1620 DGAQKEGKNKEDAKSVHFESESTSESSVLSSSSPBSTLYKEEPNGITTPGISSDVPESAH 1679
Qy      1681 KTTAASEAKSDTGOPTKVGKFOVTTTANKVGRFSVSKTEDKITDITKKEGPVAPSPFMDLEQ 1740
Db      1680 KTTAASEAKSDTGOPTKVGKFOVTTTANKVGRFSVSKTEDKITDITKKEGPVAPSPFMDLEQ 1739
Qy      1741 AVLPVAVIPKKEKPELSEPSHNLGSPDEEAAFLSKDVDGSGSPHPSQLSSKSLPSONL 1800
Db      1740 AVLPVAVIPKKEKPELSEPSHNLGSPDEEAAFLSKDVDGSGSPHPSQLSSKSLPSONL 1799
Qy      1801 SOSLSNENSSSYMSDNDSDIEDBLKLELRRLRKHKHKEIODLOSQKHIEISLYTTLG 1860
Db      1800 SOSLSNENSSSYMSDNDSDIEDBLKLELRRLRKHKHKEIODLOSQKHIEISLYTTLG 1859
Qy      1861 KVPAPAVIIPAPAPISGRRRPTKSGKSSRSSSLGNKSPOLSGNLSQSAASVYHPQOT 1920
Db      1860 KVPAPAVIIPAPAPISGRRRPTKSGKSSRSSSLGNKSPOLSGNLSQSAASVYHPQOT 1919
Qy      1921 LHPGNTPESGONLQPLKPSPSNDLYSAFTSDGAISSVLSLAPGQGSTSTTVGATV 1980

```

```

Db      1920 LHPGNTPESGONLQPLKPSPSNDLYSAFTSDGAISSVLSLAPGQGSTSTTVGATV 1979
Qy      1981 NSQAQAQPPAMTSSRKGTFTDDIHLKLVDMWARDANMLSGRRSGKGMNVEGPMARKFS 2040
Db      1980 NSQAQAQPPAMTSSRKGTFTDDIHLKLVDMWARDANMLSGRRSGKGMNVEGPMARKFS 2039
Qy      2041 APGQLCTSMTSLNGSGAPISASATSLGHFTKSCPPQOYGFPAATPFQAOMSGTGPPAQ 2100
Db      2040 APGQLCTSMTSLNGSGAPISASATSLGHFTKSCPPQOYGFPAATPFQAOMSGTGPPAQ 2099
Qy      2101 PLGQFCPVGTASLQNFNINSLQKSIINPPGSLNRTT 2136
Db      2100 PLGQFCPVGTASLQNFNINSLQKSIINPPGSLNRTT 2135

RESULT 3
ABO44384
ID ABO44384 standard; protein; 2229 AA.
XX
AC ABO44384;
XX
DT 26-SEP-2003 (first entry)
XX
DE Novel human protein kinase #4.
KW Human; kinase; enzyme; cosmetic application; nutraceutical application.
XX
OS Homo sapiens.
XX
PN US6541252-B1.
XX
PD 01-APR-2003.
XX
PF 14-MAY-2001; 2001US-00854856.
XX
PR 19-MAY-2000; 2000US-0206015P.
XX
PA (LEXI-) LEXICON GENETICS INC.
XX
PI Walke DW, Hilbun E, Donoho G, Turner CA;
XX
DR WPI; 2003-575927/54.
DR N-PSDB; ACH03767.
XX
PT New nucleic acid encoding novel human proteins, useful in cosmetic and
PS nutraceutical applications.
XX
PS Disclosure; Page; 11pp; English.
XX
CC The invention relates to a new isolated nucleic acid encoding a novel
CC human protein kinase. The nucleic acid is useful in cosmetic and
CC nutraceutical applications. The present sequence represents the amino
CC acid sequence of a novel human protein kinase. Note: The sequence data
CC for this patent did not from part of the printed specification but was
CC obtained in electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?docid=6541252B1
XX
SQ Sequence 2229 AA.

Query Match 99.5%; Score 10753.5; DB 7; Length 2229;
Best Local Similarity 95.8%; Pred. No. 0;
Matches 2136; Conservative 0; Mismatches 0; Indels 93; Gaps 1;

Qy      1 MSGGAARKQSTPSSLFLSPAPAPKXGSSSDSVGKLGAAADATYGRFEYRRRRHT 60
Db      1 MSGGAARKQSTPSSLFLSPAPAPKXGSSSDSVGKLGAAADATYGRFEYRRRRHT 60
Qy      61 MDKSRGAAATTTTTHRRFRFRSVICDSNATALELPGPLSLPPQSIIPAAVPOGAPPEPH 120
Db      61 MDKSRGAAATTTTTHRRFRFRSVICDSNATALELPGPLSLPPQSIIPAAVPOGAPPEPH 120
Qy      121 REETVTATATSGVAAQOPPAAPAAAGQAVAGPAPSTVSTSKDPVPSQPSLVSGKEBPP 180

```


Db 121 REEVTATATSQVAAQPPAAAAPEQVAPAPSTVSSISKDRPVSQPSLVSGKEPPP 180
 Qy 181 ARSSGGGSAKEPOEERSQOQDDIELEETKAVGMSNDGRFLKFDIEIGRSGFKTVYKGLD 240
 Db 181 ARSSGGGSAKEPOEERSQOQDDIELEETKAVGMSNDGRFLKFDIEIGRSGFKTVYKGLD 240
 Qy 241 TETTVAVMCELODRKLTSEORRPFKEAEMKGIQHPNIVRFYDSWESTVYKAKCIYLV 300
 Db 241 TETTVAVMCELODRKLTSEORRPFKEAEMKGIQHPNIVRFYDSWESTVYKAKCIYLV 300
 Qy 301 TELMTSGTLKTYLRPFKVMKIKVLRSWCROILKGIQPLHTPTTPIIRHDLKCDNIPTTGP 360
 Db 301 TELMTSGTLKTYLRPFKVMKIKVLRSWCROILKGIQPLHTPTTPIIRHDLKCDNIPTTGP 360
 Qy 361 TGSVYKIGDLGLATLKRAFSFAKSVIGTEPFMAPEMEYEEKYDESUVVAFGCMLEMATSEY 420
 Db 361 TGSVYKIGDLGLATLKRAFSFAKSVIGTEPFMAPEMEYEEKYDESUVVAFGCMLEMATSEY 420
 Qy 421 PYSECQNAAOIYRRVTSGVKPAKSPDKVAIPVKEIIEGCIQRONDERYSIKDILNHAFFQ 480
 Db 421 PYSECQNAAOIYRRVTSGVKPAKSPDKVAIPVKEIIEGCIQRONDERYSIKDILNHAFFQ 480
 Qy 481 EETVRVELAEEDGEXIAIKMLRIDIKKLKGKYDNDAIESFPLEDRVPEDAQEM 540
 Db 481 EETVRVELAEEDGEXIAIKMLRIDIKKLKGKYDNDAIESFPLEDRVPEDAQEM 540
 Qy 541 VESGVCYEGDHKTAKAIKDRVSLIKRRREGROLVREEOKKOEESLXQOVBQSSAQ 600
 Db 541 VESGVCYEGDHKTAKAIKDRVSLIKRRREGROLVREEOKKOEESLXQOVBQSSAQ 600
 Qy 601 TGKOLPSASTGIPLTASTSASVSTQVEPEPEADHQOLQYQOPSISVLSGTVDSGGG 660
 Db 601 TGKOLPSASTGIPLTASTSASVSTQVEPEPEADHQOLQYQOPSISVLSGTVDSGGG 660
 Qy 661 SSYFTEBSVSSQOQVSGSQHEQASHGTVPGHPISTVQAOSQPHGYPPSSVAQOGSOG 720
 Db 661 SSYFTEBSVSSQOQVSGSQHEQASHGTVPGHPISTVQAOSQPHGYPPSSVAQOGSOG 720
 Qy 721 QPSSSILTVSSSQPIQHPQOQOQIQQTAAPQQTQVYSLSTQSTSSSATTAAQAPSPQAP 780
 Db 721 QPSSSILTVSSSQPIQHPQOQOQIQQTAAPQQTQVYSLSTQSTSSSATTAAQAPSPQAP 780
 Qy 781 QVLPVVSAGKQ----- 791
 Db 781 QVLPVVSAGKQGPPEPRLPPQYPGDSNIPSSNVAVCIHSTVLXPRMPTFVLTATPGYFT 840
 Qy 792 -----STQCVSQAAPAEVAV 807
 Db 841 VNOPYVESNLVPMGVGVGOVVSQPGGSLAQAAPTSSQOAVLESTQVSVQVAPAEVAV 900
 Qy 908 AQPQATQPTTLASVDSAHSDVASGMSDGNENVPSSSGRHEGRTTKAHYKRSVRSRHE 867
 Db 908 AQPQATQPTTLASVDSAHSDVASGMSDGNENVPSSSGRHEGRTTKAHYKRSVRSRHE 867
 Qy 901 AQPQATQPTTLASVDSAHSDVASGMSDGNENVPSSSGRHEGRTTKAHYKRSVRSRHE 960
 Db 901 AQPQATQPTTLASVDSAHSDVASGMSDGNENVPSSSGRHEGRTTKAHYKRSVRSRHE 960
 Qy 968 KTSRPFKRLINAVNSKGRVVECOLETHNRKMTFKFDLDGDNPEIATIMWNNDFILAIR 927
 Db 968 KTSRPFKRLINAVNSKGRVVECOLETHNRKMTFKFDLDGDNPEIATIMWNNDFILAIR 927
 Qy 961 KTSRPFKRLINAVNSKGRVVECOLETHNRKMTFKFDLDGDNPEIATIMWNNDFILAIR 1020
 Db 961 KTSRPFKRLINAVNSKGRVVECOLETHNRKMTFKFDLDGDNPEIATIMWNNDFILAIR 1020
 Qy 928 RESFVDVREIIEKADENLSEVSVPEPSOGLESLOKQDYGPSGQKLEGEFKQPIPA 987
 Db 928 RESFVDVREIIEKADENLSEVSVPEPSOGLESLOKQDYGPSGQKLEGEFKQPIPA 987
 Qy 1021 RESFVDVREIIEKADENLSEVSVPEPSOGLESLOKQDYGPSGQKLEGEFKQPIPA 1080
 Db 1021 RESFVDVREIIEKADENLSEVSVPEPSOGLESLOKQDYGPSGQKLEGEFKQPIPA 1080
 Qy 988 SSMPOQIGIPTSSILTVVSHAGRRFIVSPVESRLRESKVPSEITDTVAASQSPGMN 1047
 Db 988 SSMPOQIGIPTSSILTVVSHAGRRFIVSPVESRLRESKVPSEITDTVAASQSPGMN 1047
 Qy 1081 SSMPOQIGIPTSSILTVVSHAGRRFIVSPVESRLRESKVPSEITDTVAASQSPGMN 1140
 Db 1081 SSMPOQIGIPTSSILTVVSHAGRRFIVSPVESRLRESKVPSEITDTVAASQSPGMN 1140
 Qy 1048 LSHSASLSLQOAFSELRRAQMTGENTAPPNFSHTGPTPVVPPLTSSIAGVPTTAAT 1107
 Db 1048 LSHSASLSLQOAFSELRRAQMTGENTAPPNFSHTGPTPVVPPLTSSIAGVPTTAAT 1107
 Qy 1141 LSHSASLSLQOAFSELRRAQMTGENTAPPNFSHTGPTPVVPPLTSSIAGVPTTAAT 1200
 Db 1141 LSHSASLSLQOAFSELRRAQMTGENTAPPNFSHTGPTPVVPPLTSSIAGVPTTAAT 1200
 Qy 1108 APVPATSSPPNDISTSVIOSEVTVPTEEGTAGVASTGVVTSGLPLTPVSESPLVSSV 1167
 Db 1108 APVPATSSPPNDISTSVIOSEVTVPTEEGTAGVASTGVVTSGLPLTPVSESPLVSSV 1167
 Qy 1201 APVPATSSPPNDISTSVIOSEVTVPTEEGTAGVASTGVVTSGLPLTPVSESPLVSSV 1260
 Db 1201 APVPATSSPPNDISTSVIOSEVTVPTEEGTAGVASTGVVTSGLPLTPVSESPLVSSV 1260

Qy 1168 SSITIPAVVSIISTSPLOVPTSTSEIVSSTALYPSTVVSATASAGGSTATPGKPPA 1227
 Db 1261 SSITIPAVVSIISTSPLOVPTSTSEIVSSTALYPSTVVSATASAGGSTATPGKPPA 1320
 Qy 1228 VVSQAAAGSTTVGATLTVSSTTSPSPSTASQLSIOLSSSTSTPTLAEVTVVSAHSLDKTS 1287
 Db 1321 VVSQAAAGSTTVGATLTVSSTTSPSPSTASQLSIOLSSSTSTPTLAEVTVVSAHSLDKTS 1380
 Qy 1288 HSSSTGLAFSLASAPSSSSSGAGVSSYISQPGGHLPIVPSVASTPTLPOAAGPTSTPL 1347
 Db 1381 HSSSTGLAFSLASAPSSSSSGAGVSSYISQPGGHLPIVPSVASTPTLPOAAGPTSTPL 1440
 Qy 1348 LPQVPSIPPLVOPANVPAVOQTLIHSGOPALBNOPHNHCPEVSDSDTOPKAGIDIDIK 1407
 Db 1441 LPQVPSIPPLVOPANVPAVOQTLIHSGOPALBNOPHNHCPEVSDSDTOPKAGIDIDIK 1500
 Qy 1408 TLEBKRLSLESHSSSGAQAASVSLSTSVIESTVTPGIPITTAVAPSKLITSTSTCLP 1467
 Db 1501 TLEBKRLSLESHSSSGAQAASVSLSTSVIESTVTPGIPITTAVAPSKLITSTSTCLP 1560
 Qy 1468 TNLPLGTVALPMTPVVTPPGQVSTPVSTTSGVKGPTAPSKPEPLTKAPVULPVGTLEPAGTL 1527
 Db 1561 TNLPLGTVALPMTPVVTPPGQVSTPVSTTSGVKGPTAPSKPEPLTKAPVULPVGTLEPAGTL 1620
 Qy 1528 PSEQLPFPFPSPSTLQSGQPLEDDAQLRSLSEKXIVTSAVGVSWAAPPTATBEAGTOP 1587
 Db 1621 PSEQLPFPFPSPSTLQSGQPLEDDAQLRSLSEKXIVTSAVGVSWAAPPTATBEAGTOP 1680
 Qy 1588 OKGVSVYKEGVPVLTATSSGACVFKMGKRPQVSAADGAQKEGKNSEDAKSVHFESESTESS 1647
 Db 1681 OKGVSVYKEGVPVLTATSSGACVFKMGKRPQVSAADGAQKEGKNSEDAKSVHFESESTESS 1740
 Qy 1648 VLSSSSPRESLTVRPEPNGITIPGISDVPESAKHTTAASEKSDTGQPTKGRFOVTTAN 1707
 Db 1741 VLSSSSPRESLTVRPEPNGITIPGISDVPESAKHTTAASEKSDTGQPTKGRFOVTTAN 1800
 Qy 1708 KYGRFVSYSKTEDKITDTPKKEGPVAPSPFMDLEQAVLPAVLPKKEKPELSPSHNGSSSD 1767
 Db 1801 KYGRFVSYSKTEDKITDTPKKEGPVAPSPFMDLEQAVLPAVLPKKEKPELSPSHNGSSSD 1860
 Qy 1768 PEAAFLSRVDVDSGSPHSPHOLSKSLPSQNLQSLSNSFNSSYMSSDNESEDIEDEDLK 1827
 Db 1861 PEAAFLSRVDVDSGSPHSPHOLSKSLPSQNLQSLSNSFNSSYMSSDNESEDIEDEDLK 1920
 Qy 1828 LEIRRLRDKHLKEIOLQOSQKHEISLYTKLGKVPNAVLIIPPAAPLSGRRRRPTYSKGS 1887
 Db 1921 LEIRRLRDKHLKEIOLQOSQKHEISLYTKLGKVPNAVLIIPPAAPLSGRRRRPTYSKGS 1980
 Qy 1888 KSSRSSSLGKNSPOLSGNLSSGQSAASVLPQOQLHPNGNIPESGQNLLOPLKPSPSDN 1947
 Db 1981 KSSRSSSLGKNSPOLSGNLSSGQSAASVLPQOQLHPNGNIPESGQNLLOPLKPSPSDN 2040
 Qy 1948 LYSALFTSDGALISVLSAPQOGTSSNTVGAATVNSQAQAQAPAMTSSSRKGTFTDDLHLK 2007
 Db 2041 LYSALFTSDGALISVLSAPQOGTSSNTVGAATVNSQAQAQAPAMTSSSRKGTFTDDLHLK 2100
 Qy 2068 GHFTKMCPPQOYGFPATPFGAQMSGTGPAPQPLQOFQVGVASIQNFISLQOSISN 2127
 Db 2161 GHFTKMCPPQOYGFPATPFGAQMSGTGPAPQPLQOFQVGVASIQNFISLQOSISN 2220
 Qy 2128 PGGSNLRTT 2136
 Db 2221 PGGSNLRTT 2229

RESULT 4
 ABP71619
 ID ABP71619 standard; protein; 2382 AA.

XX	ABF71619;
AC	29-MAY-2003 (first entry)
DT	
XX	
DE	Human MNK1 protein.
XX	
KW	MNK1; MNK4; hypotensive; gene therapy; antisense therapy; vaccine; pseudohypaldersteronism type II; PHA II; human.
XX	
OS	Homo sapiens.
XX	
PN	WO2003007793-A2.
XX	
PD	30-JAN-2003.
PP	
PF	17-JUL-2002; 2002MO-USO22744.
XX	
FR	17-JUL-2001; 2001US-0306084P.
XX	
PA	(UYVA) UNIV YALE. (LIFT/) LIFTON R P. (WILS/) WILSON F H. (CHOA/) CHOATE K (NELS/) NELSON-WILLIAMS C. (ISHI/) ISHIKAWA K.
PI	Lifton RP, Wilson FH, Choate K, Nelson-Williams C, Ishikawa K;
DR	MRI; 2003-289755/28. N-PBDB; ABZ59199.
PT	Mammalian MNK nucleic acids and proteins useful for the diagnosis, prevention and treatment of hypertension and pseudohyopaldersteronism type II.
PS	Claim 13; Page 101-109; 116pp; English.
CC	The invention relates to novel nucleic acids encoding mammalian MNK, especially human MNK1 and MNK4 proteins. The MNK polypeptides can be expressed by standard recombinant methodology. The MNK polynucleotides, polypeptides and modulators may be used in the prevention, diagnosis and treatment of diseases associated with inappropriate MNK expression such as hypertension and pseudohyopaldersteronism type II (PHA II). The present sequence represents the human MNK1 protein
SQ	Sequence 2382 AA;
Query Match	98.8%; Score 10677; DB 6; Length 2382;
Best Local Similarity	89.6%; Pred. No. 0;
Matches 2135; Conservative	0; Mismatches 1; Indels 246; Gaps 1
OY	1 MSGAAEKOSSTPGSLFLSPAPAPKNSSSDSVGEKGAAAADAVTGRTEYYRRRHT 60
Db	1 MSGGAEEKOSSTPGSLFLSPAPAPKNSSSDSVGEKGAAAAADVGRTEEYRRRHT 60
OY	61 MDKDNRGAAATTTTTEHRRFFRSVCDNSATLELPGLPLSIPOPSIPAAYQSAPPBH 120
Db	61 MDKDNRGAAATTTTTEHRRFFRSVICDSNATILELPGLPLSIPOPSIPAAVQSAPPBH 120
OY	121 REETVATATISGVAAOPPAAAAAGEQAVALGPAPSTVPSSTSKDRVSGOPSLVGKEEPPP 180
Db	121 REETVATATISGVAAOPPAAAAAGEQAVALGPAPSTVPSSTSKDRVSGOPSLVGKEEPPP 180
OY	181 ARSGGGGSASAKPEERSQQODDIELETKAAGMSNDGFELKFIDIEIGSGFKTYKKGD 240
Db	181 ARSGGGGSASAKPEERSQQODDIELETKAAGMSNDGFELKFIDIEIGSGFKTYKKGD 240
OY	241 TETTVEVMCELODKRLTSESRORFEAEEMKKGQHPIAVFYQSWESTVGXKCIVLY 300
Db	241 TETTVEVMCELQDKRLTSESRORFREBEMLKGQHPIAVFYQSWESTVGXKCIVLY 300
OY	301 TELMNGTLTKYLKRFVKMKIKVLBSWCQIIKGIQLFHTTRTPLIHRDLKDNFIITGP 360

D	b		301	TElMTSGLTKYlKRfKVMKlKYLRSWCROlLKGLOfLhTRTPPlIhRDLKCNlNlPTGP	360
Q	y		361	TGSVKIDlGhATLKRSfKSVlGfPEhMARPMEKEKDESVDVYAfGCMlEMATSEY	420
D	b		361	TGSVKIDlGhATLKRSfKSVlGfPEhMARPMEKEKDESVDVYAfGCMlEMATSEY	420
Q	y		421	PYSCQNAAOlYRRWTSGVAPSFDKVALPEVEnllEGClRONKDERYSfKDLlNHAfQ	480
D	b		421	PYSCQNAAOlYRRWTSGVAPSFDKVALPEVEnllEGClRONKDERYSfKDLlNHAfQ	480
Q	y		481	FEtSVRELAEEDDEKlAlKMLRIEDlKlKGXYKDNAlEFSPLEhVEDVDAQEM	540
D	b		481	FEtSVRELAEEDDEKlAlKMLRIEDlKlKGXYKDNAlEFSPLEhVEDVDAQEM	540
Q	y		541	VESGYVEGHhTMAKAlKORVSlfKKRfEPOPLVRECEKKQfEESlLQfVESSASQ	600
D	b		541	VESGYVEGHhTMAKAlKORVSlfKKRfEPOPLVRECEKKQfEESlLQfVESSASQ	600
Q	y		601	TGfKOLPSASTGlPTASTTSASVStQVEPEPEADhQOLQfYQOPSlSVlSDGTVDSGQ	660
D	b		601	TGfKOLPSASTGlPTASTTSASVStQVEPEPEADhQOLQfYQOPSlSVlSDGTVDSGQ	660
Q	y		661	SSVTEBSRVSSQOVASVGSQHEOASHGTVPghlPSTVVAOSOPhGYPPSSVAQGSQ	720
D	b		661	SSVTEBSRVSSQOVASVGSQHEOASHGTVPghlPSTVVAOSOPhGYPPSSVAQGSQ	720
Q	y		721	QPSSSlLTVSSSQPlGhPOQOQGlQGTAPPOQTVOYSlSQTSTSEATTAGfAPVSOPQAP	780
D	b		721	QPSSSlLTVSSSQPlGhPOQOQGlQGTAPPOQTVOYSlSQTSTSEATTAGfAPVSOPQAP	780
Q	y		781	QVLPfVNSAGKQ-----	791
D	b		781	QVLPfVNSAGKQ-----	791
Q	y		792	-----	791
D	b		841	VSQlPlSTPhVSTAGTGFSSPlPTMAGlTQPlLTLTlSAtTAlPEVSTVPSQPlTL	900
Q	y		792	-----	791
D	b		901	QPlVTLPSOVhPOLLQPAVOSMGlPAnLGOAAEVPfLSSGhVLYOGfPRLPEQYgDSNTl	960
Q	y		792	-----	791
D	b		961	APSSNVASVClhSTVfLSPMfPEVlATPgyfPfyVQfVYESNlLVPMgVGYGQVQVSQg	1020
Q	y		792	-----	834
D	b		1021	GSfLAQfPTSSQAVLSTQGSQVAAEPfAVAVQOPATOPfTLTlSSVDSASHSDVASGMS	1080
Q	y		835	DGNENVPVSSGSRHhGRTTKhRYhRSVSVSRShhEhTSPKlRlLNVSSKGPVhVCOLETh	894
D	b		1081	DGNENVPVSSGSRHhGRTTKhRYhRSVSVSRShhEhTSPKlRlLNVSSKGPVhVCOLETh	1140
Q	y		895	NRKWVTEfKFDLDGNNPEhATlVNNQfFlAlERESfVDQVREllIEKADEMLSEDSVEP	954
D	b		1141	NRKWVTEfKFDLDGNNPEhATlVNNQfFlAlERESfVDQVREllIEKADEMLSEDSVEP	1200
Q	y		955	EGDQGLESLQKNDYhGfSGOKLEGEfKQPlPASMSQOGlPTSSlTOVhVhSGRRfVl	1014
D	b		1201	EGDQGLESLQKNDYhGfSGOKLEGEfKQPlPASMSQOGlPTSSlTOVhVhSGRRfVl	1260
Q	y		1015	SPVfESRLRESKVPfPESElTDTVAASTQSPGMNlSHSASSlLQAFSELRRAQMTGEPN	1074
D	b		1261	SPVfESRLRESKVPfPESElTDTVAASTQSPGMNlSHSASSlLQAFSELRRAQMTGEPN	1320
Q	y		1075	TAPPNfSHGTfPVPVPPfLSSlAGVPTTAAAlPVPATSSPNDlSTVlQSEVYfPTE	1134
D	b		1321	TAPPNfSHGTfPVPVPPfLSSlAGVPTTAAAlPVPATSSPNDlSTVlQSEVYfPTE	1380
Q	y		1135	EGlAGVAnSTGVMtSGGlPPVfESPVfLSSVSSlTPAVVSlSTTSPSLQVfPSTSEl	1194

Db 1381 EGIAGVATSTGVTWVSGGLPIPPVESEPVLSVSVSTITIPAVVSTSTSPSLQVPTSTSEI 1440
QY 1195 VVSSALYPSYVTSATSNASAGSTATPGPKPPAVVSOQAAGSTTWTGATLTSTSTSPS 1254
Db 1441 VVSSALYPSYVTSATSNASAGSTATPGPKPPAVVSOQAAGSTTWTGATLTSTSTSPS 1500
QY 1255 TASQSLSTLSSSTSTPTLTAEFTVWVAHSLDKTSHSSTTGLAPSLAPSSSSPGAGVSSY 1314
Db 1501 TASQSLSTLSSSTSTPTLTAEFTVWVAHSLDKTSHSSTTGLAPSLAPSSSSPGAGVSSY 1560
QY 1315 ISQPGHLPVLPSTVIASTPTLPQAAGTSTPLPQVPSIPPLVQPVAVNAVQOTLIHS 1374
Db 1561 ISQPGHLPVLPSTVIASTPTLPQAAGTSTPLPQVPSIPPLVQPVAVNAVQOTLIHS 1620
QY 1375 QPOBALPNOPTHCEVSDTOPKAPGIDIKTLEELKRSIFSEHSSGAGHVSLET 1434
Db 1621 QPOBALPNOPTHCEVSDTOPKAPGIDIKTLEELKRSIFSEHSSGAGHVSLET 1680
QY 1435 SLVIESTVTGIPPTTAAVAPSKLTSTSTGCLPPTNLPIGTVALPVTVPVTPGVSTPVST 1494
Db 1681 SLVIESTVTGIPPTTAAVAPSKLTSTSTGCLPPTNLPIGTVALPVTVPVTPGVSTPVST 1740
QY 1495 TTSGVKGTAASKRPLTKAPVLPVGTLPACTLPSEQLPPPPGSLTOSQOPLDLDAOL 1554
Db 1741 TTSGVKGTAASKRPLTKAPVLPVGTLPACTLPSEQLPPPPGSLTOSQOPLDLDAOL 1800
QY 1555 RRTLSPEXITVTSVAVGPMAPATAITEAGTQPKGVQVEGCVLATSAGVFFKMGKF 1614
Db 1801 RRTLSPEXITVTSVAVGPMAPATAITEAGTQPKGVQVEGCVLATSAGVFFKMGKF 1860
QY 1615 QVSVAAADGAKGKXKSEDAKSVHEBSTSSSVLSSSSPSTLVKEPENGITTPGISSD 1674
Db 1861 QVSVAAADGAKGKXKSEDAKSVHEBSTSSSVLSSSSPSTLVKEPENGITTPGISSD 1920
QY 1675 VPESAHKTTASSEASDTPGQPTKVGPFQVTTANKVGFVSFKTDDKTITDTPKEGPVAPSP 1734
Db 1921 VPESAHKTTASSEASDTPGQPTKVGPFQVTTANKVGFVSFKTDDKTITDTPKEGPVAPSP 1980
QY 1735 FMDLEQAVLPVAVIPKKEPELSEPSHLNGPSSDPEAAFLSDVDGSGSPHSPIQLSKS 1794
Db 1981 FMDLEQAVLPVAVIPKKEPELSEPSHLNGPSSDPEAAFLSDVDGSGSPHSPIQLSKS 2040
QY 1795 LPSQNLQSLSNSFNSSWSSDNESEDIEDDLKLELRRLRDXHKEIQDLQSRQKHIES 1854
Db 2041 LPSQNLQSLSNSFNSSWSSDNESEDIEDDLKLELRRLRDXHKEIQDLQSRQKHIES 2100
QY 1855 LYTLGKVPVPAVILPPAAPLSGRRRRTKSGSSSSSSSGNKSPLQSGNLSCQSAASY 1914
Db 2101 LYTLGKVPVPAVILPPAAPLSGRRRRTKSGSSSSSSSGNKSPLQSGNLSCQSAASY 2160
QY 1915 LHPQQTLPHPGNIPESSQNLQPLKSPSSSDNLYSAFTSDGATSVPSLSAPGGTSTTN 1974
Db 2161 LHPQQTLPHPGNIPESSQNLQPLKSPSSSDNLYSAFTSDGATSVPSLSAPGGTSTTN 2220
QY 1975 TVGATVNSQAQAQPPAMTSSRKGTFTDDLHKLVDNMAPDAMNLGSRGSGKHNNYEGPG 2034
Db 2221 TVGATVNSQAQAQPPAMTSSRKGTFTDDLHKLVDNMAPDAMNLGSRGSGKHNNYEGPG 2280
QY 2035 MARFESAPEGULCISMTSNLGGSAPIISAASATSLGHFTKSMCPPOQYFPATPFAQMSGT 2094
Db 2281 MARFESAPEGULCISMTSNLGGSAPIISAASATSLGHFTKSMCPPOQYFPATPFAQMSGT 2340
QY 2095 GGPAQPLGQFPQVGTASLQNFNISNLQKSIINPPGSLNRTT 2136
Db 2341 GGPAQPLGQFPQVGTASLQNFNISNLQKSIINPPGSLNRTT 2382

RESULT 5
ABO44381
ID ABO44381 standard; protein; 2382 AA.
XX AC
XX ABO44381;
XX

DT 26-SEP-2003 (first entry)
XX
DE Novel human protein kinase #1.
XX
KW Human; kinase; enzyme; cosmetic application; nutraceutical application.
XX
OS Homo sapiens.
XX
PN US6541252-B1.
XX
PD 01-APR-2003.
XX
PF 14-MAY-2001; 2001US-00854856.
XX
PR 19-MAY-2000; 2000US-020601SP.
XX
PA (TEXT-) LEXICON GENETICS INC.
XX
PI Walke DW, Hilbun E, Donoho G, Turner CA;
XX
DR WPI: 2003-575927/54.
XX
DR N-PSDB: ACH03764.
XX
PT New nucleic acid encoding novel human proteins, useful in cosmetic and
PT nutraceutical applications.
PS
SS
XX
XX Disclosure; Page: 11pp; English.
CC The invention relates to a new isolated nucleic acid encoding a novel
CC human protein kinase. The nucleic acid is useful in cosmetic and
CC nutraceutical applications. The present sequence represents the amino
CC acid sequence of a novel human protein kinase. Note: The sequence data
CC for this patent did not from part of the printed specification but was
CC obtained in electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?docid=6541252B1
CC
XX
SQ Sequence 2382 AA;
Query Match 98.8%; Score 10677; DB 7; Length 2382;
Best Local Similarity 89.7%; Pred. No. 0;
Matches 2136; Conservative 0; Mismatches 0; Indels 246; Gaps 1;
QY 1 MSGGAAGKQSTGSLFLSPAPAPKNGSSDSSVGEKLGAAADAVTGRTTEYRRRHT 60
Db 1 MSGGAAGKQSTGSLFLSPAPAPKNGSSDSSVGEKLGAAADAVTGRTTEYRRRHT 60
QY 61 MDXDSRGAATTTTTEHFRFRRSVTCDSNATALELPPLPSLPQSPIDPAVPGSAPPEPH 120
Db 61 MDXDSRGAATTTTTEHFRFRRSVTCDSNATALELPPLPSLPQSPIDPAVPGSAPPEPH 120
QY 121 REETVATATSOVAQCPAAAGGQAVAGPASTVSSSTSKORPVSPISVSGKEPP 180
Db 121 REETVATATSOVAQCPAAAGGQAVAGPASTVSSSTSKORPVSPISVSGKEPP 180
QY 181 ARSGSGGSAKEPOEERSQOQDDIELEETRAVAGSNDGRFLKFDIEIGRGSFKTVYKGLD 240
Db 181 ARSGSGGSAKEPOEERSQOQDDIELEETRAVAGSNDGRFLKFDIEIGRGSFKTVYKGLD 240
QY 241 TETTVEAVACHELDRKLTKEERORFKEEAMKGLQHPNIVRFYDSWESTVKGKCIIVL 300
Db 241 TETTVEAVACHELDRKLTKEERORFKEEAMKGLQHPNIVRFYDSWESTVKGKCIIVL 300
QY 301 TELMTSGTLKTYLKRFFVMKIKVLRSGCRQIILGLOLHRTPTPIIHRDLKCNIFITGP 360
Db 301 TELMTSGTLKTYLKRFFVMKIKVLRSGCRQIILGLOLHRTPTPIIHRDLKCNIFITGP 360
QY 361 TGSVKIGDLGLATLKASFAKSVIGTPEFMAPEMVEKXDESVDVAVFGCMLEMATSEY 420
Db 361 TGSVKIGDLGLATLKASFAKSVIGTPEFMAPEMVEKXDESVDVAVFGCMLEMATSEY 420
QY 421 PYSECQMAAQIYRRVYSGVPASFDKVAIPEVKEIIEGCIKQKDERYSIKDLINHAFFQ 480
Db 421 PYSECQMAAQIYRRVYSGVPASFDKVAIPEVKEIIEGCIKQKDERYSIKDLINHAFFQ 480

```

QY 481 EETGVRELEAEDEGEKAIKLMIRIEDIKLKGGYKDNFAIESFDLERDVEDVAQEM 540
Db 481 EETGVRELEAEDEGEKAIKLMIRIEDIKLKGGYKDNFAIESFDLERDVEDVAQEM 540
QY 541 VESGVYEGDGHKHTAKAIKDRVSLIKRRKEQRLVREHEKKKOBESSLKOQVQSSASQ 600
Db 541 VESGVYEGDGHKHTAKAIKDRVSLIKRRKEQRLVREHEKKKOBESSLKOQVQSSASQ 600
QY 601 TGIKOLPSASTGIPASTTSASVSTQVREPEEADQHOOLOYOQPSISVLSDGTVDSGQG 660
Db 601 TGIKOLPSASTGIPASTTSASVSTQVREPEEADQHOOLOYOQPSISVLSDGTVDSGQG 660
QY 661 SSVTESSVSSQOQVSVSGSHOAHSTGTVPNGH PSTQAQSOQPHGVPPSSVAQGSQSG 720
Db 661 SSVTESSVSSQOQVSVSGSHOAHSTGTVPNGH PSTQAQSOQPHGVPPSSVAQGSQSG 720
QY 721 QPSSSSILTVGSSSQPIQHPPQOQOQIGQCTAPPOQTVQVSLSTQSSSEATTAAQVSOQAP 780
Db 721 QPSSSSILTVGSSSQPIQHPPQOQOQIGQCTAPPOQTVQVSLSTQSSSEATTAAQVSOQAP 780
QY 781 QVLPQVSAAGKO----- 791
Db 781 QVLPQVSAAGKO----- 791
QY 792 ----- 791
Db 792 ----- 791
QY 901 QPVTQLPSQVHPQLLPQAVQSMGIPANLGAQAEVPLSGVDLYOGPPRLPPQVPGDSNI 960
Db 901 QPVTQLPSQVHPQLLPQAVQSMGIPANLGAQAEVPLSGVDLYOGPPRLPPQVPGDSNI 960
QY 792 ----- 791
Db 792 ----- 791
QY 961 APSNNVASVCIHSTVLXRPMPTEVLATPGYPTVQVYVBSNLLVPMGVGQVQVQSGPQ 1020
Db 961 APSNNVASVCIHSTVLXRPMPTEVLATPGYPTVQVYVBSNLLVPMGVGQVQVQSGPQ 1020
QY 792 ----- 834
Db 792 ----- 834
QY 1021 GSLAQAFTTSSQOQVLESTQOQVAPAEVAVAAQOQPTTLASSVDSASHSDVYASMS 1080
Db 1021 GSLAQAFTTSSQOQVLESTQOQVAPAEVAVAAQOQPTTLASSVDSASHSDVYASMS 1080
QY 835 DGNENVPSSSGRHEGRTTKRHYRKSVRSSRHEKTSREKRLILNVSNKGRDVECOLETH 894
Db 835 DGNENVPSSSGRHEGRTTKRHYRKSVRSSRHEKTSREKRLILNVSNKGRDVECOLETH 894
QY 1081 DGNENVPSSSGRHEGRTTKRHYRKSVRSSRHEKTSREKRLILNVSNKGRDVECOLETH 1140
Db 1081 DGNENVPSSSGRHEGRTTKRHYRKSVRSSRHEKTSREKRLILNVSNKGRDVECOLETH 1140
QY 895 NRKAVTFFDLQDGNPEEATILIMVNDPILAIERSFPVDQVREIIEKADMLSEDSYEP 954
Db 895 NRKAVTFFDLQDGNPEEATILIMVNDPILAIERSFPVDQVREIIEKADMLSEDSYEP 954
QY 1141 NRKAVTFFDLQDGNPEEATILIMVNDPILAIERSFPVDQVREIIEKADMLSEDSYEP 1200
Db 1141 NRKAVTFFDLQDGNPEEATILIMVNDPILAIERSFPVDQVREIIEKADMLSEDSYEP 1200
QY 955 EGDQGLESLQKDDYGFSGSQKLEGEFKOPIPASMPQOIGIPTSSLTQVYVHSAGRPIY 1014
Db 955 EGDQGLESLQKDDYGFSGSQKLEGEFKOPIPASMPQOIGIPTSSLTQVYVHSAGRPIY 1014
QY 1201 EGDQGLESLQKDDYGFSGSQKLEGEFKOPIPASMPQOIGIPTSSLTQVYVHSAGRPIY 1260
Db 1201 EGDQGLESLQKDDYGFSGSQKLEGEFKOPIPASMPQOIGIPTSSLTQVYVHSAGRPIY 1260
QY 1015 SPVESRRLRESKVPSEIITDTVAASSTAQSPGMLSHSASLSLQOAFSELRAAQTBEPN 1074
Db 1015 SPVESRRLRESKVPSEIITDTVAASSTAQSPGMLSHSASLSLQOAFSELRAAQTBEPN 1074
QY 1261 SPVESRRLRESKVPSEIITDTVAASSTAQSPGMLSHSASLSLQOAFSELRAAQTBEPN 1320
Db 1261 SPVESRRLRESKVPSEIITDTVAASSTAQSPGMLSHSASLSLQOAFSELRAAQTBEPN 1320
QY 1075 TAPRPFSGTGTFFVPPFLSSIAGVPTTAATAAPVAPATSSPPNDISTSVYQSEVTVYTE 1134
Db 1075 TAPRPFSGTGTFFVPPFLSSIAGVPTTAATAAPVAPATSSPPNDISTSVYQSEVTVYTE 1134
QY 1321 TAPRPFSGTGTFFVPPFLSSIAGVPTTAATAAPVAPATSSPPNDISTSVYQSEVTVYTE 1380
Db 1321 TAPRPFSGTGTFFVPPFLSSIAGVPTTAATAAPVAPATSSPPNDISTSVYQSEVTVYTE 1380
QY 1135 EGIAGVATSTGVTVSGGLPIPVSESPVLSVSSITIPAVVSIITSPSIQVPTSISEI 1194
Db 1135 EGIAGVATSTGVTVSGGLPIPVSESPVLSVSSITIPAVVSIITSPSIQVPTSISEI 1194
QY 1381 EGIAGVATSTGVTVSGGLPIPVSESPVLSVSSITIPAVVSIITSPSIQVPTSISEI 1440
Db 1381 EGIAGVATSTGVTVSGGLPIPVSESPVLSVSSITIPAVVSIITSPSIQVPTSISEI 1440
QY 1195 VVSSTALVPSVTVATSAAGSTATPGKPPAVVSOQAAGSTTVGATLTSVSTTSPS 1254
Db 1195 VVSSTALVPSVTVATSAAGSTATPGKPPAVVSOQAAGSTTVGATLTSVSTTSPS 1254
QY 1441 VVSSTALVPSVTVATSAAGSTATPGKPPAVVSOQAAGSTTVGATLTSVSTTSPS 1500
Db 1441 VVSSTALVPSVTVATSAAGSTATPGKPPAVVSOQAAGSTTVGATLTSVSTTSPS 1500
QY 1255 TASQLSTQLSSSTSTPTLAEVTVVSAHSLDXTSHSTTGLAFSLSABSSSSPPAGVSSY 1314
Db 1255 TASQLSTQLSSSTSTPTLAEVTVVSAHSLDXTSHSTTGLAFSLSABSSSSPPAGVSSY 1314
QY 1501 TASQLSTQLSSSTSTPTLAEVTVVSAHSLDXTSHSTTGLAFSLSABSSSSPPAGVSSY 1560
Db 1501 TASQLSTQLSSSTSTPTLAEVTVVSAHSLDXTSHSTTGLAFSLSABSSSSPPAGVSSY 1560

```

```

QY 1315 ISQPGHLPLVTPSVIASTPILPOAAGPTSTPLLPOVPSIPLVQVYANVAVOQTLLHS 1374
Db 1315 ISQPGHLPLVTPSVIASTPILPOAAGPTSTPLLPOVPSIPLVQVYANVAVOQTLLHS 1374
QY 1375 QCPALLPNOPTHCPREVSDTPQKABGIDDIKLLEKRLSLFSEHSSSGAQAHSVLET 1434
Db 1375 QCPALLPNOPTHCPREVSDTPQKABGIDDIKLLEKRLSLFSEHSSSGAQAHSVLET 1434
QY 1621 QCPALLPNOPTHCPREVSDTPQKABGIDDIKLLEKRLSLFSEHSSSGAQAHSVLET 1680
Db 1621 QCPALLPNOPTHCPREVSDTPQKABGIDDIKLLEKRLSLFSEHSSSGAQAHSVLET 1680
QY 1435 SLVIESVTTPGIPPTTANVAPSKLTLSTTSCLPPTNPLGLTVALLVTPVTVTGQVSTPST 1494
Db 1435 SLVIESVTTPGIPPTTANVAPSKLTLSTTSCLPPTNPLGLTVALLVTPVTVTGQVSTPST 1494
QY 1681 SLVIESVTTPGIPPTTANVAPSKLTLSTTSCLPPTNPLGLTVALLVTPVTVTGQVSTPST 1740
Db 1681 SLVIESVTTPGIPPTTANVAPSKLTLSTTSCLPPTNPLGLTVALLVTPVTVTGQVSTPST 1740
QY 1495 TTSQVXKGTAPSKRPLTKAPVLPVGTLPAGTLPESEQLPPEPGSLTQSOQPLDLDLQQL 1554
Db 1495 TTSQVXKGTAPSKRPLTKAPVLPVGTLPAGTLPESEQLPPEPGSLTQSOQPLDLDLQQL 1554
QY 1741 TTSQVXKGTAPSKRPLTKAPVLPVGTLPAGTLPESEQLPPEPGSLTQSOQPLDLDLQQL 1800
Db 1741 TTSQVXKGTAPSKRPLTKAPVLPVGTLPAGTLPESEQLPPEPGSLTQSOQPLDLDLQQL 1800
QY 1555 RRLTSPXKITVTSANVGVSMAPPAITTEAGTQPKGVSVQKEGVLATSSGAGVFKMGRE 1614
Db 1555 RRLTSPXKITVTSANVGVSMAPPAITTEAGTQPKGVSVQKEGVLATSSGAGVFKMGRE 1614
QY 1801 RRLTSPXKITVTSANVGVSMAPPAITTEAGTQPKGVSVQKEGVLATSSGAGVFKMGRE 1860
Db 1801 RRLTSPXKITVTSANVGVSMAPPAITTEAGTQPKGVSVQKEGVLATSSGAGVFKMGRE 1860
QY 1615 QVSVAAADGAQKEGKNKSEDAKSVFEESTSTSSVLSSSPESTLVKPEPNTITPGISSD 1674
Db 1615 QVSVAAADGAQKEGKNKSEDAKSVFEESTSTSSVLSSSPESTLVKPEPNTITPGISSD 1674
QY 1861 QVSVAAADGAQKEGKNKSEDAKSVFEESTSTSSVLSSSPESTLVKPEPNTITPGISSD 1920
Db 1861 QVSVAAADGAQKEGKNKSEDAKSVFEESTSTSSVLSSSPESTLVKPEPNTITPGISSD 1920
QY 1675 VPESAHKTTASSEAKSDTGOPTKVGRFOVTTTANKVGRFSYKTEDKITDTKEGPVVASPP 1734
Db 1675 VPESAHKTTASSEAKSDTGOPTKVGRFOVTTTANKVGRFSYKTEDKITDTKEGPVVASPP 1734
QY 1921 VPESAHKTTASSEAKSDTGOPTKVGRFOVTTTANKVGRFSYKTEDKITDTKEGPVVASPP 1980
Db 1921 VPESAHKTTASSEAKSDTGOPTKVGRFOVTTTANKVGRFSYKTEDKITDTKEGPVVASPP 1980
QY 1735 FMDLEQAVLPAVITPKKEKPELSEPSHLNGPSDPEAAFLSHVDVDDGSGSPHSPIQLSSKS 1794
Db 1735 FMDLEQAVLPAVITPKKEKPELSEPSHLNGPSDPEAAFLSHVDVDDGSGSPHSPIQLSSKS 1794
QY 1981 FMDLEQAVLPAVITPKKEKPELSEPSHLNGPSDPEAAFLSHVDVDDGSGSPHSPIQLSSKS 2040
Db 1981 FMDLEQAVLPAVITPKKEKPELSEPSHLNGPSDPEAAFLSHVDVDDGSGSPHSPIQLSSKS 2040
QY 1795 LPSQNLQSLSSNSNSNSNSNSNSNSNSNSNSNSNSNSNSNSNSNSNSNSNSNSNSNSNS 1854
Db 1795 LPSQNLQSLSSNSNSNSNSNSNSNSNSNSNSNSNSNSNSNSNSNSNSNSNSNSNSNSNS 1854
QY 2041 LPSQNLQSLSSNSNSNSNSNSNSNSNSNSNSNSNSNSNSNSNSNSNSNSNSNSNSNSNS 2100
Db 2041 LPSQNLQSLSSNSNSNSNSNSNSNSNSNSNSNSNSNSNSNSNSNSNSNSNSNSNSNSNS 2100
QY 1855 LYTGLGVPPPAVITPPAAPLSGRRRRPTKSKGSSRSSSLGNKSPOLSGNLSGQSAASV 1914
Db 1855 LYTGLGVPPPAVITPPAAPLSGRRRRPTKSKGSSRSSSLGNKSPOLSGNLSGQSAASV 1914
QY 2101 LYTGLGVPPPAVITPPAAPLSGRRRRPTKSKGSSRSSSLGNKSPOLSGNLSGQSAASV 2160
Db 2101 LYTGLGVPPPAVITPPAAPLSGRRRRPTKSKGSSRSSSLGNKSPOLSGNLSGQSAASV 2160
QY 1915 LHPQOTLHPPGNIPESQONOLQPLKPSPSSDNLYSAAFTSDGAISVPSLSAPGQGTSTTN 1974
Db 1915 LHPQOTLHPPGNIPESQONOLQPLKPSPSSDNLYSAAFTSDGAISVPSLSAPGQGTSTTN 1974
QY 2161 LHPQOTLHPPGNIPESQONOLQPLKPSPSSDNLYSAAFTSDGAISVPSLSAPGQGTSTTN 2220
Db 2161 LHPQOTLHPPGNIPESQONOLQPLKPSPSSDNLYSAAFTSDGAISVPSLSAPGQGTSTTN 2220
QY 1975 TVGATVNSQAAQOPPAMTSSRKGTFTDDLHKLVDMNARDAMNLSGRGSGHNNYBGP 2034
Db 1975 TVGATVNSQAAQOPPAMTSSRKGTFTDDLHKLVDMNARDAMNLSGRGSGHNNYBGP 2034
QY 2221 TVGATVNSQAAQOPPAMTSSRKGTFTDDLHKLVDMNARDAMNLSGRGSGHNNYBGP 2280
Db 2221 TVGATVNSQAAQOPPAMTSSRKGTFTDDLHKLVDMNARDAMNLSGRGSGHNNYBGP 2280
QY 2035 MARKFSAPGOLCISMTSNLGGSAPIISAASATSLGHFTKSMCPPOQYGFPAATPGQAWSGT 2094
Db 2035 MARKFSAPGOLCISMTSNLGGSAPIISAASATSLGHFTKSMCPPOQYGFPAATPGQAWSGT 2094
QY 2281 MARKFSAPGOLCISMTSNLGGSAPIISAASATSLGHFTKSMCPPOQYGFPAATPGQAWSGT 2340
Db 2281 MARKFSAPGOLCISMTSNLGGSAPIISAASATSLGHFTKSMCPPOQYGFPAATPGQAWSGT 2340
QY 2095 GGPAPQPLGQFQPVGTASLQNFNLSNLOKSI SNPPGSLRTT 2136
Db 2095 GGPAPQPLGQFQPVGTASLQNFNLSNLOKSI SNPPGSLRTT 2136
QY 2341 GGPAPQPLGQFQPVGTASLQNFNLSNLOKSI SNPPGSLRTT 2382
Db 2341 GGPAPQPLGQFQPVGTASLQNFNLSNLOKSI SNPPGSLRTT 2382

```

RESULT 6
ABO4395
ID ABO4395 standard; protein; 2108 AA.

ABO4395;
XX
DT 26-SEP-2003 (first entry)

XX Novel human protein kinase #15.
XX Human; kinase; enzyme; cosmetic application; nutraceutical application.
XX Homo sapiens.
XX OS
XX PN US6541252-B1.

01-APR-2003.
 14-MAY-2001; 2001US-00854856.
 19-MAY-2000; 2000US-0206015P.
 (LEXI-) LEXICON GENETICS INC.
 Walke DW, Hilbun E, Donoho G, Turner CA;
 MPI; 2003-575927/54.
 N-PSDB; ACH03778.
 New nucleic acid encoding novel human proteins, useful in cosmetic and
 nutritional applications.
 Disclosure; Page; 11pp; English.
 The invention relates to a new isolated nucleic acid encoding a novel
 human protein kinase. The nucleic acid is useful in cosmetic and
 nutritional applications. The present sequence represents the amino
 acid sequence of a novel human protein kinase. Note: The sequence data
 for this patent did not from part of the printed specification but was
 obtained in electronic format directly from USPTO at
 seqdata.uspto.gov/sequence.html?docid=6541252B1
 Sequence 2108 AA;
 Query Match 98.5%; Score 10646; DB 7; Length 2108;
 Best Local Similarity 98.7%; Pred. No. 0;
 Matches 2108; Conservative 0; Mismatches 0; Indels 28; Gaps 1;
 1 MSGGAAGKOSTPESLFLSPAPAPAKNGSSSDSVGKLGAAADAVTGRTEERYRRRT 60
 1 MSGGAAGKOSTPESLFLSPAPAPAKNGSSSDSVGKLGAAADAVTGRTEERYRRRT 60
 61 MDKDSGCAATTTTHERFRFRSVTCDNATALELPGLPLSLPQPSIPAAVPSAPPEPH 120
 61 MDKDSGCAATTTTHERFRFRSVTCDNATALELPGLPLSLPQPSIPAAVPSAPPEPH 120
 121 REETVATATSOVAQOPPAAPAEQAVAPAPSTVSSSKORPVQPSLVSKKEPP 180
 121 REETVATATSOVAQOPPAAPAEQAVAPAPSTVSSSKORPVQPSLVSKKEPP 180
 121 REETVATATSOVAQOPPAAPAEQAVAPAPSTVSSSKORPVQPSLVSKKEPP 180
 181 ARSGSGGSAKEPEERSQODDIELEETKAVGMSNDGRFLKEDIBIGRSFKTVYKGLD 240
 181 ARSGSGGSAKEPEERSQODDIELEETKAVGMSNDGRFLKEDIBIGRSFKTVYKGLD 240
 241 TETTVEAVACELDRKLTKESEORFKEAEMLKGLQHPNI VRFYDSESTVKGKCI VLV 300
 241 TETTVEAVACELDRKLTKESEORFKEAEMLKGLQHPNI VRFYDSESTVKGKCI VLV 300
 301 TELMTSGTLKTYLKRPFVMKIXYLRSWCROILKGLQHLHRTPIIHRDLKCNIFITGP 360
 301 TELMTSGTLKTYLKRPFVMKIXYLRSWCROILKGLQHLHRTPIIHRDLKCNIFITGP 360
 361 TGSVKIGDLGLATLKRAFSAKSVIGTPEFMAPEMEYEKEDSVYVAFGKMLEMATSEY 420
 361 TGSVKIGDLGLATLKRAFSAKSVIGTPEFMAPEMEYEKEDSVYVAFGKMLEMATSEY 420
 421 PYSECQNAQIYRRVTSVGVKPAFDKVAIPVEKKEIEGCI RQNKDERYSIKDLINAFQ 480
 421 PYSECQNAQIYRRVTSVGVKPAFDKVAIPVEKKEIEGCI RQNKDERYSIKDLINAFQ 480
 481 EETGVAVELAEEDDGERIAKIMLRITEDIKGLKGYKDNDAIFSPFLLEDVVEDVAQEM 540
 481 EETGVAVELAEEDDGERIAKIMLRITEDIKGLKGYKDNDAIFSPFLLEDVVEDVAQEM 540
 541 VESGVYCEGDHKTMAKAIKDRVSLIKRRKQROLVREEOEKKEQESLSKQVEQSSASQ 600
 541 VESGVYCEGDHKTMAKAIKDRVSLIKRRKQROLVREEOEKKEQESLSKQVEQSSASQ 600
 601 TGIKOLPSASTGIPTASTTSASVSTOVEPEPEADHQOLQYOQPSISVLSDGTVDGSG 660

601 TGIKOLPSASTGIPTASTTSASVSTOVEPEPEADHQOLQYOQPSISVLSDGTVDGSG 660
 661 SSVFTESSVSSQOQVSGSQHEQASHGTGPHGIPSTVQAQSOQPHGYPRESSVAQSGSQ 720
 661 SSVFTESSVSSQOQVSGSQHEQASHGTGPHGIPSTVQAQSOQPHGYPRESSV----- 713
 721 QPSSSLTGVSSTQPIQHPOQOQGIQOTAPPOQTVQSLSGTSTSTSTSTSTSTSTSTSTST 780
 714 -----QQGIQOTAPPOQTVQSLSGTSTSTSTSTSTSTSTSTSTSTSTSTSTSTST 752
 781 QVLPQVSAKOSTQGVSAQVAPAEVVAQOQATPPTLASSVDSVSHSDVSAQMGDGENV 840
 753 QVLPQVSAKOSTQGVSAQVAPAEVVAQOQATPPTLASSVDSVSHSDVSAQMGDGENV 812
 841 PSSSGRHEGRTTKRHKYSVRSNRHEKTSRPKRIINVSNGKDRVVECOLETHNRKVT 900
 813 PSSSGRHEGRTTKRHKYSVRSNRHEKTSRPKRIINVSNGKDRVVECOLETHNRKVT 872
 901 FKFDLDNDEBEIATIMVNDPILATRESFVQVREIIRKADMLSEDSVVEBEGQGL 960
 873 FKFDLDNDEBEIATIMVNDPILATRESFVQVREIIRKADMLSEDSVVEBEGQGL 932
 961 ESLQKQDYGFSQSKLEGEFKQPIPASSMPQOIGIPTSSLTVYHSAGRFIVSPVES 1020
 933 ESLQKQDYGFSQSKLEGEFKQPIPASSMPQOIGIPTSSLTVYHSAGRFIVSPVES 992
 1021 RLRESKVFSEITDYVAASTAQSPGMNLSHSAASLSIQAFSELRRQOMTEGPTAPNF 1080
 993 RLRESKVFSEITDYVAASTAQSPGMNLSHSAASLSIQAFSELRRQOMTEGPTAPNF 1052
 1081 SHTGPTFPVVPFLSSIAGVPTAAATAPPAVATSSPNDISTSVTQSEVVPVEBGIAGV 1140
 1053 SHTGPTFPVVPFLSSIAGVPTAAATAPPAVATSSPNDISTSVTQSEVVPVEBGIAGV 1112
 1141 ATSTGVVTSGLPIPVSESPVLSVVSITIPAVVISITPSSLOQPTSTSEIVSSTA 1200
 1113 ATSTGVVTSGLPIPVSESPVLSVVSITIPAVVISITPSSLOQPTSTSEIVSSTA 1172
 1201 LPSVTVSATSAAGSSTATPGRKPAVVSQAAGSTVTCALLTVSTTSPSTASQLS 1260
 1173 LPSVTVSATSAAGSSTATPGRKPAVVSQAAGSTVTCALLTVSTTSPSTASQLS 1232
 1261 IQSSSTPTTAEVTVVASHSLDKTSHSSTTGLASLSAPSSSSSGAGVSSYISQPG 1320
 1233 IQSSSTPTTAEVTVVASHSLDKTSHSSTTGLASLSAPSSSSSGAGVSSYISQPG 1292
 1321 LHEPLVPSVIASTPILPOAAGPTSTPLLPQVPSIPPLVQPVANVAVOQTLHISQOPAL 1380
 1293 LHEPLVPSVIASTPILPOAAGPTSTPLLPQVPSIPPLVQPVANVAVOQTLHISQOPAL 1352
 1381 LPHQPHTHCPENDSDPKAPGIDDKITLBEKRLSFSEHSSGAAQHASVLETSVIES 1440
 1353 LPHQPHTHCPENDSDPKAPGIDDKITLBEKRLSFSEHSSGAAQHASVLETSVIES 1412
 1441 TVTPGIPPTAVASKLLTSTSTCLPPTNLPLGTVALPTPVYTPQGVSPVSTTSGVK 1500
 1413 TVTPGIPPTAVASKLLTSTSTCLPPTNLPLGTVALPTPVYTPQGVSPVSTTSGVK 1472
 1501 PGTAPEKPLTKAPVLPVGTETLPGATLPSBOLPPFPGPSLITOSQQLLEDLDALRRTLSP 1560
 1473 PGTAPEKPLTKAPVLPVGTETLPGATLPSBOLPPFPGPSLITOSQQLLEDLDALRRTLSP 1532
 1533 EXITTVSAVGPVMAAPTAITEAGTPOKGVSVKGPVLAISSGAVFMGRFQVSVAA 1592
 1561 EXITTVSAVGPVMAAPTAITEAGTPOKGVSVKGPVLAISSGAVFMGRFQVSVAA 1620
 1621 DGAQKGGKKSBDKAKVHFESSSTSSSVSSSPESSTLVYPEBNGATTIPGISDVESAH 1680
 1593 DGAQKGGKKSBDKAKVHFESSSTSSSVSSSPESSTLVYPEBNGATTIPGISDVESAH 1652
 1681 KTTASAKSDTQOPTVGRFQVTTTANKVGRFVSKTEDIKITTKKEGPVAVSPFMDLQ 1740

Db 1653 KTTASAKSDTGTGQTKYGRFOVTTANKGRFVSXKTEDEK1TDTKKEGVASPEFMDLEQ 1712
 QY 1741 AVLPAYLPKKEKPELSPSHLNGSPSPPEAFISRDVDDSGSPHSHQISSKSLPEQNTL 1800
 Db 1713 AVLPAYLPKKEKPELSPSHLNGSPSPPEAFISRDVDDSGSPHSHQISSKSLPEQNTL 1772
 QY 1801 SOSLSNFSNYSNDSNEDSDEDDLELKLRLDKLKEIODLQSQKHEIESLYTKLG 1860
 Db 1773 SOSLSNFSNYSNDSNEDSDEDDLELKLRLDKLKEIODLQSQKHEIESLYTKLG 1832
 QY 1861 KVPAPVILPPAPPLSGRRRPTKSGKSSRSSSLGNKSPQLSGNLSGQSAASYLHPQQT 1920
 Db 1833 KVPAPVILPPAPPLSGRRRPTKSGKSSRSSSLGNKSPQLSGNLSGQSAASYLHPQQT 1892
 QY 1921 LHPPGNIPESGQNLQPLKPSPESSDNLKSAFTSDGAI SVPSLSAPQGGISSTTVGATV 1980
 Db 1893 LHPPGNIPESGQNLQPLKPSPESSDNLKSAFTSDGAI SVPSLSAPQGGISSTTVGATV 1952
 QY 1981 NSQAAQAPAMTSSRGKFTFDLHKLVDMWARMAMLSGRGSGKGMNVEGPMARKFS 2040
 Db 1953 NSQAAQAPAMTSSRGKFTFDLHKLVDMWARMAMLSGRGSGKGMNVEGPMARKFS 2012
 QY 2041 APQGLCSMTSNLGGSAPIASAATSLGHFTKSNCPQYGFPAFPAGQWSTGAPAPQ 2100
 Db 2013 APQGLCSMTSNLGGSAPIASAATSLGHFTKSNCPQYGFPAFPAGQWSTGAPAPQ 2072
 QY 2101 PLGQFQVGTASLQENFISNLQKSI SNPPGSNLFTT 2136
 Db 2073 PLGQFQVGTASLQENFISNLQKSI SNPPGSNLFTT 2108

RESULT 7
 ADC99049
 ID ADC99049 standard; protein: 2108 AA.
 AC ADC99049;
 DT 01-JAN-2004 (first entry)
 DE Human KRP protein - SEQ ID 2.
 XX anti-HIV; antiallergic; antiinflammatory; antianaemic; antiparkinsonian;
 KW neotropic; anticonvulsant; antiarteriosclerotic; antiaschemtic;
 KW immunosuppressive; antihypertoid; cytoskeletal; hepatotropic; dermatological;
 KW antidiabetic; nephrotoxic; antitumor; thyromimetic; neuroprotective;
 KW osteopathic; antiparasitic; antipneumonic; antihelminthic; antiparasitic;
 KW uteropathic; ophthalmological; antihypertensive; haemostatic; antibacterial;
 KW virucide; protozoacide; fungicide; kinase; phosphatase; KRP;
 KW cell proliferative disorder; atherosclerosis; cirrhosis; hepatitis;
 KW cancer; developmental; mental retardation; neurological;
 KW Alzheimer's disease; Parkinson's; autoimmune; inflammatory; Crohn's;
 KW diabetes mellitus; viral; bacterial; fungal; parasitic; protozoan;
 KW helminthic infection; transgenic; gene therapy; human; enzyme.
 OS Homo sapiens.
 PN WO2003033680-A2.
 PD 24-APR-2003.
 PF 17-OCT-2002; 2002WO-US033723.
 PR 19-OCT-2001; 2001US-0345474P.
 PR 02-NOV-2001; 2001US-0343910P.
 PR 13-NOV-2001; 2001US-0333098P.
 PR 16-NOV-2001; 2001US-0332424P.
 PR 30-NOV-2001; 2001US-0334288P.
 PA (INCY-) INCYTE GENOMICS INC.
 PI Bandman O, Baughn MR, Becha SD, Borowsky MT, Dugan BM,
 PI Emerling BM, Forsythe IU, Gandhi AR, Gorvad AE, Griffin JA,
 PI Gururajan R, Hafalia AJA, Khan FA, Lal PG, Lee EA, Lee SY;

PI Lindquist EA, Lu DM, Lu Y, Marguis JP, Nguyen DB, Arvizu CS;
 PI Ramkumar J, Recipon SA, Richardson TW, Swarnakar A, Tang YT;
 PI Thornton MB, Iran UK, Chawla NK, Warren BA, Yang U, Yao MG, Yue H;
 PI Zebayadian Y;
 DR WPI; 2003-403214/38.
 DR N-PSDB; ADC99101.
 XX
 PT New human kinases and phosphatases and polynucleotides, useful for
 PT diagnosing, treating or preventing autoimmune or inflammatory disorders
 PT (e.g. AIDS, allergy or anemia), multiple sclerosis, osteoarthritis,
 PT cancer or hepatitis.
 PS Claim 1; SEQ ID NO 2; 424pp; English.
 CC
 CC The invention relates to a novel isolated polypeptide which is a human
 CC kinase and phosphatase (KRP). The KRP polypeptides, polynucleotides,
 CC agonists and antagonists are useful for diagnosing, treating or
 CC preventing cell proliferative disorders such as atherosclerosis,
 CC cirrhosis, hepatitis and cancer, developmental disorders e.g. mental
 CC retardation, neurological disorders including Alzheimer's disease and
 CC Parkinson's disease, autoimmune and inflammatory disorders such as
 CC Crohn's disease and diabetes mellitus and finally, viral, bacterial,
 CC fungal, parasitic, protozoan or helminthic infections. Furthermore, the
 CC polynucleotides encoding KRP may be useful for creating transgenic
 CC animals to model human disease, as well as during gene therapy
 CC procedures. The current sequence is that of the human KRP protein of the
 CC invention.
 CC
 CC
 SQ Sequence 2108 AA;
 Query Match 98.5%; Score 10646; DB 7; Length 2108;
 Best Local Similarity 98.6%; Pred. No. 0;
 Matches 2107; Conservative 0; Mismatches 1; Indels 28; Gaps 1;
 QY 1 MSGGAEEKOSSTPGSLPLSPAPAPKXSSSDSSVGEKLGAAADAVGTETEEYRRRHT 60
 Db 1 MSGGAEEKOSSTPGSLPLSPAPAPKXSSSDSSVGEKLGAAADAVGTETEEYRRRHT 60
 QY 61 MDKDSRGAAATTTTTERFRFRSVICSNATALELPPLSLPQPSIPAAVPSAPPEPH 120
 Db 61 MDKDSRGAAATTTTTERFRFRSVICSNATALELPPLSLPQPSIPAAVPSAPPEPH 120
 QY 121 REETVTATSOVAQOPPAAPAEQAVAPASTVSSSTKDPVQPSLVGSKPEPP 180
 Db 121 REETVTATSOVAQOPPAAPAEQAVAPASTVSSSTKDPVQPSLVGSKPEPP 180
 QY 181 ARSGSGGSAKEPQEBRSQODDIELEETKAVGMSNDGRELKPDIEIGRSFKTVYGLD 240
 Db 181 ARSGSGGSAKEPQEBRSQODDIELEETKAVGMSNDGRELKPDIEIGRSFKTVYGLD 240
 QY 241 TETTVEAVKELQDRKTKSEBQPFKEAEMLKGLOHPNIVRFYDSMESTYKGGKCLIV 300
 Db 241 TETTVEAVKELQDRKTKSEBQPFKEAEMLKGLOHPNIVRFYDSMESTYKGGKCLIV 300
 QY 301 TELMTSGTLKTYLKRPFVKMKIYLRSWCROILKGLQFLHRTPEIHRDLKCDNIFITGP 360
 Db 301 TELMTSGTLKTYLKRPFVKMKIYLRSWCROILKGLQFLHRTPEIHRDLKCDNIFITGP 360
 QY 361 TGSVKIGDLGLATLKRAAFKSVIETPEFMAPEMYEERKYDESVDVYAFGCMLEMATSEY 420
 Db 361 TGSVKIGDLGLATLKRAAFKSVIETPEFMAPEMYEERKYDESVDVYAFGCMLEMATSEY 420
 QY 421 PYSECQNAAOIYRVRYSVGVKPSADKAIPEVVEITIGCRONDEYYSIKDLINHAFFQ 480
 Db 421 PYSECQNAAOIYRVRYSVGVKPSADKAIPEVVEITIGCRONDEYYSIKDLINHAFFQ 480
 QY 481 EETGVRELAEBDDGEKIAIKMLRIDIKLKKYKQNDNAIESFLERDVEDVAQEM 540
 Db 481 EETGVRELAEBDDGEKIAIKMLRIDIKLKKYKQNDNAIESFLERDVEDVAQEM 540
 QY 541 VESGYVGDHKTMAKAIKDRVSLIKKKEORQLVREBQKKQESSLKQOVQSSASQ 600
 Db 541 VESGYVGDHKTMAKAIKDRVSLIKKKEORQLVREBQKKQESSLKQOVQSSASQ 600

Db 541 VESGIVCEGDKHTMAKAIKDRVSLIKRRKREQLRVREOEKKOEESLKQVQESASQ 600
 QY 601 TGIKQLPEASTGIPATSTTSASVSTQVEPEPEADQHQOLQYQOQPSISVLSDGTVDSSQG 660
 Db 601 TGIKQLPEASTGIPATSTTSASVSTQVEPEPEADQHQOLQYQOQPSISVLSDGTVDSSQG 660
 QY 661 SSVPTESSVSSQOQTVSSYSGOHQAHSTGTVPQH1PSTVQAOQOPHGVYPPSSVAQOQSG 720
 Db 661 SSVPTESSVSSQOQTVSSYSGOHQAHSTGTVPQH1PSTVQAOQOPHGVYPPSSVAQOQSG 720
 QY 721 QPSSSSLTGVSSSQIQHPQOQOQGIQCTAPPOQTVQVLSQTSSTSSSEKTTKQPVSQOPAP 780
 Db 721 QPSSSSLTGVSSSQIQHPQOQOQGIQCTAPPOQTVQVLSQTSSTSSSEKTTKQPVSQOPAP 780
 QY 781 QVLQVQVSAKQOSTQGVQVAPAEVAVNAQOPATOPTTASSVDASHSVASGMDGNV 840
 Db 781 QVLQVQVSAKQOSTQGVQVAPAEVAVNAQOPATOPTTASSVDASHSVASGMDGNV 840
 QY 841 PSSSGRHEGRTTKHRYKRSVRSRHEKTSRPKLRLNVSNKGRVVECOLETHNRKMT 900
 Db 841 PSSSGRHEGRTTKHRYKRSVRSRHEKTSRPKLRLNVSNKGRVVECOLETHNRKMT 900
 QY 901 FKFLDGNPEEIKATIMVNDPILAIERESFVDQVREIIEKADMLSDVSVVEEGDGL 960
 Db 901 FKFLDGNPEEIKATIMVNDPILAIERESFVDQVREIIEKADMLSDVSVVEEGDGL 960
 QY 961 ESLOGKDDYGSQKLEGEFKOP1PASSMPQOIGIPSSLTQVHVSAGRFIYSPVES 1020
 Db 961 ESLOGKDDYGSQKLEGEFKOP1PASSMPQOIGIPSSLTQVHVSAGRFIYSPVES 1020
 QY 993 ESLOGKDDYGSQKLEGEFKOP1PASSMPQOIGIPSSLTQVHVSAGRFIYSPVES 992
 Db 993 ESLOGKDDYGSQKLEGEFKOP1PASSMPQOIGIPSSLTQVHVSAGRFIYSPVES 992
 QY 1021 RLRESKVPSEITDTVAASTQSPQMLSHSASSLSLQOAFSELRRAQMTGEPYAPBNF 1080
 Db 1021 RLRESKVPSEITDTVAASTQSPQMLSHSASSLSLQOAFSELRRAQMTGEPYAPBNF 1080
 QY 993 RLRESKVPSEITDTVAASTQSPQMLSHSASSLSLQOAFSELRRAQMTGEPYAPBNF 1052
 Db 993 RLRESKVPSEITDTVAASTQSPQMLSHSASSLSLQOAFSELRRAQMTGEPYAPBNF 1052
 QY 1081 SHTGPTPVVPPFSLISLNGVTTAAATAPVPATSSPPNDISTVYQSEVYTPTEBGLAGV 1140
 Db 1081 SHTGPTPVVPPFSLISLNGVTTAAATAPVPATSSPPNDISTVYQSEVYTPTEBGLAGV 1140
 QY 1053 SHTGPTPVVPPFSLISLNGVTTAAATAPVPATSSPPNDISTVYQSEVYTPTEBGLAGV 1112
 Db 1053 SHTGPTPVVPPFSLISLNGVTTAAATAPVPATSSPPNDISTVYQSEVYTPTEBGLAGV 1112
 QY 1141 ATSTGVVTSAGGLPIPPVESPVLSVSVSITIPAVVSIITSPSLQVPTSEIIVSSTA 1200
 Db 1141 ATSTGVVTSAGGLPIPPVESPVLSVSVSITIPAVVSIITSPSLQVPTSEIIVSSTA 1200
 QY 1113 ATSTGVVTSAGGLPIPPVESPVLSVSVSITIPAVVSIITSPSLQVPTSEIIVSSTA 1172
 Db 1113 ATSTGVVTSAGGLPIPPVESPVLSVSVSITIPAVVSIITSPSLQVPTSEIIVSSTA 1172
 QY 1201 LYPVTVSATSASAGSTATPGKPPAVVSOQAAGSTTVGATLTVSITTSFPPSTAQSL 1260
 Db 1201 LYPVTVSATSASAGSTATPGKPPAVVSOQAAGSTTVGATLTVSITTSFPPSTAQSL 1260
 QY 1173 LYPVTVSATSASAGSTATPGKPPAVVSOQAAGSTTVGATLTVSITTSFPPSTAQSL 1232
 Db 1173 LYPVTVSATSASAGSTATPGKPPAVVSOQAAGSTTVGATLTVSITTSFPPSTAQSL 1232
 QY 1261 IQLSSSTPTLAEVTVVSAHSLDKTSHSSTTGLAFSLAPSSSSSPGAGVSYISOPGG 1320
 Db 1261 IQLSSSTPTLAEVTVVSAHSLDKTSHSSTTGLAFSLAPSSSSSPGAGVSYISOPGG 1320
 QY 1233 IQLSSSTPTLAEVTVVSAHSLDKTSHSSTTGLAFSLAPSSSSSPGAGVSYISOPGG 1292
 Db 1233 IQLSSSTPTLAEVTVVSAHSLDKTSHSSTTGLAFSLAPSSSSSPGAGVSYISOPGG 1292
 QY 1321 LHPVTVSVIASTPTLPQAGPTSTPLPOVPSIPPLVQVAVVAVOQOTLIHSOPAPAL 1380
 Db 1321 LHPVTVSVIASTPTLPQAGPTSTPLPOVPSIPPLVQVAVVAVOQOTLIHSOPAPAL 1380
 QY 1293 LHPVTVSVIASTPTLPQAGPTSTPLPOVPSIPPLVQVAVVAVOQOTLIHSOPAPAL 1352
 Db 1293 LHPVTVSVIASTPTLPQAGPTSTPLPOVPSIPPLVQVAVVAVOQOTLIHSOPAPAL 1352
 QY 1361 LPPNPHTHCPEVDSPTQKAPGIDIDIKTLEKLSLFESEHSSSGAQAHSVLSLETSLVIES 1440
 Db 1361 LPPNPHTHCPEVDSPTQKAPGIDIDIKTLEKLSLFESEHSSSGAQAHSVLSLETSLVIES 1440
 QY 1353 LPPNPHTHCPEVDSPTQKAPGIDIDIKTLEKLSLFESEHSSSGAQAHSVLSLETSLVIES 1412
 Db 1353 LPPNPHTHCPEVDSPTQKAPGIDIDIKTLEKLSLFESEHSSSGAQAHSVLSLETSLVIES 1412
 QY 1441 TVTPTGIPPTTAAVAPSKLTLSTSTCLPPTNLPLGTALVTVTVTVPGVSTVSTTSGVK 1500
 Db 1441 TVTPTGIPPTTAAVAPSKLTLSTSTCLPPTNLPLGTALVTVTVTVPGVSTVSTTSGVK 1500
 QY 1501 PGTPAPSKPEPLTKAPVLPVGTLPAGTLPSEQLPPPGSLTQSQOPLDLQAKRLTISP 1560
 Db 1501 PGTPAPSKPEPLTKAPVLPVGTLPAGTLPSEQLPPPGSLTQSQOPLDLQAKRLTISP 1560
 QY 1473 PGTPAPSKPEPLTKAPVLPVGTLPAGTLPSEQLPPPGSLTQSQOPLDLQAKRLTISP 1532
 Db 1473 PGTPAPSKPEPLTKAPVLPVGTLPAGTLPSEQLPPPGSLTQSQOPLDLQAKRLTISP 1532
 QY 1561 EXITVTSVAVGVSMAPTAITEAGTQPKGVQVVEGCVLATSAGVFKMGRFQVSVAA 1620
 Db 1561 EXITVTSVAVGVSMAPTAITEAGTQPKGVQVVEGCVLATSAGVFKMGRFQVSVAA 1620
 QY 1533 EITVTSVAVGVSMAPTAITEAGTQPKGVQVVEGCVLATSAGVFKMGRFQVSVAA 1592
 Db 1533 EITVTSVAVGVSMAPTAITEAGTQPKGVQVVEGCVLATSAGVFKMGRFQVSVAA 1592
 QY 1621 DGAQEGKCKSEADASVHFESESTSSSVLSSSSPESTLVKPEPNKITPGISSDVPEBAH 1680
 Db 1621 DGAQEGKCKSEADASVHFESESTSSSVLSSSSPESTLVKPEPNKITPGISSDVPEBAH 1680
 QY 1593 DGAQEGKCKSEADASVHFESESTSSSVLSSSSPESTLVKPEPNKITPGISSDVPEBAH 1652
 Db 1593 DGAQEGKCKSEADASVHFESESTSSSVLSSSSPESTLVKPEPNKITPGISSDVPEBAH 1652

QY 1681 KTTASEAKSDPTGQFTKVGREQVTTTANKVGRFSVSKTEDKITDITKKGFPVAPPPMDLEQ 1740
 Db 1681 KTTASEAKSDPTGQFTKVGREQVTTTANKVGRFSVSKTEDKITDITKKGFPVAPPPMDLEQ 1740
 QY 1553 KTTASEAKSDPTGQFTKVGREQVTTTANKVGRFSVSKTEDKITDITKKGFPVAPPPMDLEQ 1712
 Db 1553 KTTASEAKSDPTGQFTKVGREQVTTTANKVGRFSVSKTEDKITDITKKGFPVAPPPMDLEQ 1712
 QY 1741 AVLPVAVPKKEKPELSPSHLNGPSSDPEEAFSLRDVDDGSGHSPHQLSSKSLPSONL 1800
 Db 1741 AVLPVAVPKKEKPELSPSHLNGPSSDPEEAFSLRDVDDGSGHSPHQLSSKSLPSONL 1800
 QY 1713 AVLPVAVPKKEKPELSPSHLNGPSSDPEEAFSLRDVDDGSGHSPHQLSSKSLPSONL 1772
 Db 1713 AVLPVAVPKKEKPELSPSHLNGPSSDPEEAFSLRDVDDGSGHSPHQLSSKSLPSONL 1772
 QY 1801 SOSLSNFSNYSMSNDSESDIEDLKLRLRLDKHLKEIQLDLSRQKHEIESLYTKLG 1860
 Db 1801 SOSLSNFSNYSMSNDSESDIEDLKLRLRLDKHLKEIQLDLSRQKHEIESLYTKLG 1860
 QY 1773 SOSLSNFSNYSMSNDSESDIEDLKLRLRLDKHLKEIQLDLSRQKHEIESLYTKLG 1832
 Db 1773 SOSLSNFSNYSMSNDSESDIEDLKLRLRLDKHLKEIQLDLSRQKHEIESLYTKLG 1832
 QY 1861 KVPVAVIIPPAAPLSGRRRRTKSGKSSSRSSSLGNKSPOLSGNLSGQSAASVJLHPQOT 1920
 Db 1861 KVPVAVIIPPAAPLSGRRRRTKSGKSSSRSSSLGNKSPOLSGNLSGQSAASVJLHPQOT 1920
 QY 1833 KVPVAVIIPPAAPLSGRRRRTKSGKSSSRSSSLGNKSPOLSGNLSGQSAASVJLHPQOT 1892
 Db 1833 KVPVAVIIPPAAPLSGRRRRTKSGKSSSRSSSLGNKSPOLSGNLSGQSAASVJLHPQOT 1892
 QY 1921 LHPGNTPEGONOLLOPLKPSPSDNLVSAFTSDGAI SVPSLSAPGQGISSTTVGATV 1980
 Db 1921 LHPGNTPEGONOLLOPLKPSPSDNLVSAFTSDGAI SVPSLSAPGQGISSTTVGATV 1980
 QY 1893 LHPGNTPEGONOLLOPLKPSPSDNLVSAFTSDGAI SVPSLSAPGQGISSTTVGATV 1952
 Db 1893 LHPGNTPEGONOLLOPLKPSPSDNLVSAFTSDGAI SVPSLSAPGQGISSTTVGATV 1952
 QY 1981 NSQAQAOPPAMTSSRRKGTFTDHLKLVDMWAPAMLSGRRSGKHMYNTEPGMARKFS 2040
 Db 1981 NSQAQAOPPAMTSSRRKGTFTDHLKLVDMWAPAMLSGRRSGKHMYNTEPGMARKFS 2040
 QY 1953 NSQAQAOPPAMTSSRRKGTFTDHLKLVDMWAPAMLSGRRSGKHMYNTEPGMARKFS 2012
 Db 1953 NSQAQAOPPAMTSSRRKGTFTDHLKLVDMWAPAMLSGRRSGKHMYNTEPGMARKFS 2012
 QY 2041 APGLCISMTSNLGGAPISASATSLGHFTKSKCPQOYGFPAATPGACWSGTGEPAPQ 2100
 Db 2041 APGLCISMTSNLGGAPISASATSLGHFTKSKCPQOYGFPAATPGACWSGTGEPAPQ 2100
 QY 2013 APGLCISMTSNLGGAPISASATSLGHFTKSKCPQOYGFPAATPGACWSGTGEPAPQ 2072
 Db 2013 APGLCISMTSNLGGAPISASATSLGHFTKSKCPQOYGFPAATPGACWSGTGEPAPQ 2072
 QY 2101 PLGQFQVGTASLQNFNINISNLOKISINPPGNSLRTT 2136
 Db 2101 PLGQFQVGTASLQNFNINISNLOKISINPPGNSLRTT 2108
 QY 2073 PLGQFQVGTASLQNFNINISNLOKISINPPGNSLRTT 2108
 Db 2073 PLGQFQVGTASLQNFNINISNLOKISINPPGNSLRTT 2108

RESULT 8
 ABO4392
 ID ABO4392 standard; protein; 2201 AA.
 XX
 AC ABO4392;
 XX
 DT 26-SEP-2003 (first entry)
 XX
 DE Novel human protein kinase #12.
 XX
 KM Human; kinase; enzyme; cosmetic application; nutraceutical application.
 XX
 OS Homo sapiens.
 XX
 PN US6541252-B1.
 XX
 PD 01-APR-2003.
 XX
 PE 14-MAY-2001; 2001US-00854856.
 XX
 PR 19-MAY-2000; 2000US-0206015P.
 XX
 PA (LEXI-) LEXICON GENETICS INC.
 XX
 PI Walke DW, Hilbun E, Donoho G, Turner CA;
 DR WP1; 2003-575927/54.
 DR N-PSDB; ACH03775.
 XX
 PT New nucleic acid encoding novel human proteins, useful in cosmetic and
 PT nutraceutical applications.
 XX
 PS Disclosure; Page; 11pp; English.
 XX
 CC The invention relates to a new isolated nucleic acid encoding a novel
 CC human protein kinase. The nucleic acid is useful in cosmetic and
 CC nutraceutical applications. The present sequence represents the amino
 CC acid sequence of a novel human protein kinase. Note: The sequence data
 CC for this patent did not from part of the printed specification but was

CC obtained in electronic format directly from USPTO at
 CC seqdata.uspto.gov/sequence.html?docid=6541252B1
 XX
 SQ Sequence 2201 AA;

Query Match 97.9%; Score 10589.5; DB 7; Length 2201;
 Best Local Similarity 94.6%; Pred. No. 0;
 Matches 2108; Conservative 0; Mismatches 0; Indels 121; Gaps 2;

QY 1 MSGAAEKOSSSTPGLFLSPAPAPKNGSSSDSSVGEKLGAAADAVVGRTEYRRRHT 60
 Db 1 MSGAAEKOSSSTPGLFLSPAPAPKNGSSSDSSVGEKLGAAADAVVGRTEYRRRHT 60
 QY 61 MDKSRGAATTTTTEHFRFRSVICDSNATALEPLGLPLSPPSIDPAVPOSAPPEPH 120
 Db 61 MDKSRGAATTTTTEHFRFRSVICDSNATALEPLGLPLSPPSIDPAVPOSAPPEPH 120
 QY 121 REETVATATSOVAAQPPMAAAGPQAVAGPAPSTVPSSTSKDRVSOPLVGSKEEPP 180
 Db 121 REETVATATSOVAAQPPMAAAGPQAVAGPAPSTVPSSTSKDRVSOPLVGSKEEPP 180
 QY 181 ARSGGGSAAKEPOERSQODDIELETKAVGMSDGRFLKFDIEIGRSGFKTYKGLD 240
 Db 181 ARSGGGSAAKEPOERSQODDIELETKAVGMSDGRFLKFDIEIGRSGFKTYKGLD 240
 QY 241 TETTVAVAMELODRKLTCSRORFKEAEMLKGLQHPNIVRFYDSWESTVKGKCIYLV 300
 Db 241 TETTVAVAMELODRKLTCSRORFKEAEMLKGLQHPNIVRFYDSWESTVKGKCIYLV 300
 QY 301 TELMTSGTLKTYLKRFKYMKIKVLRSCROILKGLQFLHTPTPIIHRDLKCDNIFITGP 360
 Db 301 TELMTSGTLKTYLKRFKYMKIKVLRSCROILKGLQFLHTPTPIIHRDLKCDNIFITGP 360
 QY 361 TGSVYIGDLGATLKRAFAKSVIGTPEFMAPEWEEKYDESVMYARGMCLMATESEY 420
 Db 361 TGSVYIGDLGATLKRAFAKSVIGTPEFMAPEWEEKYDESVMYARGMCLMATESEY 420
 QY 421 PYSBCONAAQIYRRVTSVGPASFPDKVAIPYKELIEGICIRONKDERYSIKDLINHAFFQ 480
 Db 421 PYSBCONAAQIYRRVTSVGPASFPDKVAIPYKELIEGICIRONKDERYSIKDLINHAFFQ 480
 QY 481 EETGVAVELAEEDGEEKIAIKMLRIEDIKLKKGKYKONEAIBESFEDLERDVEDVADEM 540
 Db 481 EETGVAVELAEEDGEEKIAIKMLRIEDIKLKKGKYKONEAIBESFEDLERDVEDVADEM 540
 QY 541 VESGVCEGDEKHTAKAIKDRVSLIKRREOROLVREOEKKKOBESSLKQOVEOSSASQ 600
 Db 541 VESGVCEGDEKHTAKAIKDRVSLIKRREOROLVREOEKKKOBESSLKQOVEOSSASQ 600
 QY 601 TGIKOLPASSTGIPASTTSASVSTOVPPEPEADQHQLOQYQPSISVLSGTVDSGQG 660
 Db 601 TGIKOLPASSTGIPASTTSASVSTOVPPEPEADQHQLOQYQPSISVLSGTVDSGQG 660
 QY 661 SSVTESSVSSQOQVSYSGOHEOAHSTGTVGHIPTSTQAOOSOPHGVYPPSSVAAQSGQG 720
 Db 661 SSVTESSVSSQOQVSYSGOHEOAHSTGTVGHIPTSTQAOOSOPHGVYPPSSVAAQSGQG 720
 QY 721 QPSSSILTVSSSQPIQHPQOQOGIOOTAPPOQTVOYSLSTQTSSEATTAQVPSOPAP 780
 Db 721 QPSSSILTVSSSQPIQHPQOQOGIOOTAPPOQTVOYSLSTQTSSEATTAQVPSOPAP 780
 QY 781 QVLPQVSAKGKQ-----QGGIQQATPPOQTVOYSLSTQTSSEATTAQVPSOPAP 791
 Db 781 QVLPQVSAKGKQ-----QGGIQQATPPOQTVOYSLSTQTSSEATTAQVPSOPAP 791
 QY 792 -----STQVSOVAPAEVAV 807
 Db 792 -----STQVSOVAPAEVAV 807
 QY 808 AOPATOPTTLASVDSHSDVASGMSDGENENPSSSGRHRGRTTKKHYSRVSRSRHE 867
 Db 808 AOPATOPTTLASVDSHSDVASGMSDGENENPSSSGRHRGRTTKKHYSRVSRSRHE 867
 QY 873 AOPATOPTTLASVDSHSDVASGMSDGENENPSSSGRHRGRTTKKHYSRVSRSRHE 932
 Db 873 AOPATOPTTLASVDSHSDVASGMSDGENENPSSSGRHRGRTTKKHYSRVSRSRHE 932

QY 868 KTSRPKRLILNVSNKGDREVECOLETHNRKMYTFEFDLDGDNPEIATIMVNNDFILAE 927
 Db 933 KTSRPKRLILNVSNKGDREVECOLETHNRKMYTFEFDLDGDNPEIATIMVNNDFILAE 992
 QY 928 RESFVDQVREIIEKADEMPLSEDVSEPEBDQGLESLQCKDDYGFSGSKLEGEFKOPIPA 987
 Db 993 RESFVDQVREIIEKADEMPLSEDVSEPEBDQGLESLQCKDDYGFSGSKLEGEFKOPIPA 1052
 QY 988 SSMPOQIGITPSSILTOVSHSARRRIVSPVESRRESKVPSEITTDVVAASTAOSPGMN 1047
 Db 1053 SSMPOQIGITPSSILTOVSHSARRRIVSPVESRRESKVPSEITTDVVAASTAOSPGMN 1112
 QY 1048 LSHSASSLSLQOARSELRAQMTGEPNTAPNFSHTGTFVVPVPELSIAGVPTTAAT 1107
 Db 1113 LSHSASSLSLQOARSELRAQMTGEPNTAPNFSHTGTFVVPVPELSIAGVPTTAAT 1172
 QY 1108 APVPATSSPPNDISTSVIOSEVTVPTBEGIAGVATSTGVTSVGGIPIPVSESPVLSVV 1167
 Db 1173 APVPATSSPPNDISTSVIOSEVTVPTBEGIAGVATSTGVTSVGGIPIPVSESPVLSVV 1232
 QY 1168 SSITIPAVVSIISTPSLOVPTSEIYVSTALYPSVTVSATASAGSTATGPKRPA 1227
 Db 1233 SSITIPAVVSIISTPSLOVPTSEIYVSTALYPSVTVSATASAGSTATGPKRPA 1292
 QY 1228 VVSQOAAGSTTVGATLINSVSTTSFPSTASQLSLQSSSTPTPLAETVVVSAHSLDKTS 1287
 Db 1293 VVSQOAAGSTTVGATLINSVSTTSFPSTASQLSLQSSSTPTPLAETVVVSAHSLDKTS 1352
 QY 1288 HSSTTGAFLSAPSSSSPGAGVSVYSIQPGIHLPIVPSVIASTPILPOAAPTSTPL 1347
 Db 1353 HSSTTGAFLSAPSSSSPGAGVSVYSIQPGIHLPIVPSVIASTPILPOAAPTSTPL 1412
 QY 1348 LPQVPSIPPLVQPVANPAVQOULIHQOPQALLPNOPTHCPEDVSDTOPKAGIDIK 1407
 Db 1413 LPQVPSIPPLVQPVANPAVQOULIHQOPQALLPNOPTHCPEDVSDTOPKAGIDIK 1472
 QY 1408 TLEBKRLSEFHSSSGAGHVASLETSIVESTVTFGTPTTAAVPSKLTSTSTCPLP 1467
 Db 1473 TLEBKRLSEFHSSSGAGHVASLETSIVESTVTFGTPTTAAVPSKLTSTSTCPLP 1532
 QY 1468 TNLPLGTVALPVPVTPGVSTVSTTSQVKGRTAPSKPPLTKAPVLPVGTBLPAGTL 1527
 Db 1533 TNLPLGTVALPVPVTPGVSTVSTTSQVKGRTAPSKPPLTKAPVLPVGTBLPAGTL 1592
 QY 1528 PSEOLPPFPGPSILQSOQPLEDLDAOLRRLTSPKXIVTSAVGVSAAPALITBAGOP 1587
 Db 1593 PSEOLPPFPGPSILQSOQPLEDLDAOLRRLTSPKXIVTSAVGVSAAPALITBAGOP 1652
 QY 1588 QKGVSYQKEGVFLATSSGAGVFKMGREFQVSVAAADGAKGEGNKSEDAKSVFEESTSSSS 1647
 Db 1653 QKGVSYQKEGVFLATSSGAGVFKMGREFQVSVAAADGAKGEGNKSEDAKSVFEESTSSSS 1712
 QY 1648 VLSSSPESTLYKKEPNIGITIPGISDVPSAHKTTTSEAKSDTGOPTKVGREFVYTTAN 1707
 Db 1713 VLSSSPESTLYKKEPNIGITIPGISDVPSAHKTTTSEAKSDTGOPTKVGREFVYTTAN 1772
 QY 1708 KVGFRSVSKTEDKTTDPKKGCPVVASPPFMDLEQAVLPAVLPKKEPELSEESHNGPSSD 1767
 Db 1773 KVGFRSVSKTEDKTTDPKKGCPVVASPPFMDLEQAVLPAVLPKKEPELSEESHNGPSSD 1832
 QY 1768 PEAFLSRDYDDGSGSPHSPHQLSSKSLPSQNSQSLSNSNSNSYMSDNESEDIEDLX 1827
 Db 1833 PEAFLSRDYDDGSGSPHSPHQLSSKSLPSQNSQSLSNSNSNSYMSDNESEDIEDLX 1892
 QY 1828 LELARRLDKHLKEIQQDIQSORKEHIESLYTKLGVPVPAVILPPAAPLSGRRRRPTSKGS 1887
 Db 1893 LELARRLDKHLKEIQQDIQSORKEHIESLYTKLGVPVPAVILPPAAPLSGRRRRPTSKGS 1952
 QY 1888 KSSRSSSLGNKSPOLSGNLSGQSAASYLHPQOULHPGNIPESSQONULLQPLKSPSSDN 1947
 Db 1953 KSSRSSSLGNKSPOLSGNLSGQSAASYLHPQOULHPGNIPESSQONULLQPLKSPSSDN 2012
 QY 1948 LYSAFITSDGAISVPSLAPGQSTSTVTVGATVNSQAQAQPPAMTSSRKGTFTDIDLHL 2007

|||||
Db 2013 LYSAFISDGAISVSLSPAGQGTSTINVGATVNSQAAPPMATSSRKCTFTDDHLKL 2072
Oy 2008 VDNARDDMNLSGRSGSKGHNNVGGPGMARKPSAPGOLCISMTSNLGGASAPISASATSL 2067
Db 2073 VDNARDDMNLSGRSGSKGHNNVGGPGMARKPSAPGOLCISMTSNLGGASAPISASATSL 2132
Oy 2068 GHFTKSMCPPOQYGFPAFPAGQWSGTGGAPQPLGQFPQVGTASLQNFNISNLQKSTISN 2127
Db 2133 GHFTKSMCPPOQYGFPAFPAGQWSGTGGAPQPLGQFPQVGTASLQNFNISNLQKSTISN 2192
Oy 2128 PPGSNLRTT 2136
Db 2193 PPGSNLRTT 2201

RESULT 9
ABO44389
ID ABO44389 standard; protein, 2354 AA.
XX ABO44389;
AC ABO44389;
XX 26-SEP-2003 (first entry)
DT
XX Novel human protein kinase #9.
DE
XX Human; kinase; enzyme; cosmetic application; nutraceutical application.
KW Homo sapiens.
XX
OS Homo sapiens.
XX
PN US6541252-B1.
PD 01-APR-2003.
XX
PE 14-MAY-2001; 2001US-00854856.
XX
PR 19-MAY-2000; 2000US-0206015P.
XX
PA (LEXI-) LEXICON GENETICS INC.
XX
FI Walke DW, Hilbun E, Donoho G, Turner CA;
XX
DR WPI; 2003-575927/54.
DR N-PSDB; ACH03772.
XX
PT New nucleic acid encoding novel human proteins, useful in cosmetic and
XX nutraceutical applications.
XX
PS Disclosure; Page; 11pp; English.
XX
CC The invention relates to a new isolated nucleic acid encoding a novel
CC human protein kinase. The nucleic acid is useful in cosmetic and
CC nutraceutical applications. The present sequence represents the amino
CC acid sequence of a novel human protein kinase. Note: The sequence data
CC for this patent did not from part of the printed specification but was
CC obtained in electronic format directly from USPTO at
CC segdata.uspto.gov/sequence.html?docid=6541252B1
XX
SQ Sequence 2354 AA;

Query Match 97.2%; Score 10513; DB 7; Length 2354;
Best Local Similarity 88.5%; Pred. No. 0;
Matches 2108; Conservative 0; Mismatches 0; Indels 274; Gaps 2;

Oy 1 MSGAAAEQOSTPGSLFLSPAPAPKNGSSSDSVGEKLGAAADAVTGRTEFYRRRHT 60
Db 1 MSGAAAEQOSTPGSLFLSPAPAPKNGSSSDSVGEKLGAAADAVTGRTEFYRRRHT 60
Oy 61 MDKDSRGAATTTTTEHFRFRSVICDSNATALELPGILPLSPPSIPAAVQSAAPPBH 120
Db 61 MDKDSRGAATTTTTEHFRFRSVICDSNATALELPGILPLSPPSIPAAVQSAAPPBH 120
Oy 121 REETVATATSOVAQPPAAAAAGBOAVAGPAPSTVPSSTSKDREVPVQSIVGSKXEPBP 180

|||||
Db 121 REETVATATSOVAQPPAAAAAGBOAVAGPAPSTVPSSTSKDREVPVQSIVGSKXEPBP 180
Oy 181 ARSGGGGSAKEPEEESQOQDDIELEETAYAVGNSNGRFLKFIETIGRGSFKTYKGLD 240
Db 181 ARSGGGGSAKEPEEESQOQDDIELEETAYAVGNSNGRFLKFIETIGRGSFKTYKGLD 240
Oy 241 TETVEAVAMELODRKLTKEERCFKEAEMLKGLQHPNIVRFDSWSESTYKGGKCIYLV 300
Db 241 TETVEAVAMELODRKLTKEERCFKEAEMLKGLQHPNIVRFDSWSESTYKGGKCIYLV 300
Oy 301 TELMTSGTLKTYLKRFFVMKIKVLRSWCRQILKGLQPLHRTTPIIHRDLKCDNIPTIGP 360
Db 301 TELMTSGTLKTYLKRFFVMKIKVLRSWCRQILKGLQPLHRTTPIIHRDLKCDNIPTIGP 360
Oy 361 TGSVKIGDIGLATLKRASFASVGTPEEPNAPMEYEEKXDSVVVAFGCMLEMAVSEY 420
Db 361 TGSVKIGDIGLATLKRASFASVGTPEEPNAPMEYEEKXDSVVVAFGCMLEMAVSEY 420
Oy 421 PYSECQNAAQIYRVTSGVKPASFDKVAIPVKEIIIEGCIKQNDERSIKDLNHAFFQ 480
Db 421 PYSECQNAAQIYRVTSGVKPASFDKVAIPVKEIIIEGCIKQNDERSIKDLNHAFFQ 480
Oy 481 EETGVAVELAEEDDEGEKIALKMLRIDIKKLGKTYKQNEALIESFDLEDPVPEVQEM 540
Db 481 EETGVAVELAEEDDEGEKIALKMLRIDIKKLGKTYKQNEALIESFDLEDPVPEVQEM 540
Oy 541 VESGVYCEGCHKTKAKXIKRVSILKRREORQVREBOEKKQESSLKQOVQSSASQ 600
Db 541 VESGVYCEGCHKTKAKXIKRVSILKRREORQVREBOEKKQESSLKQOVQSSASQ 600
Oy 601 TGIKQPSASTGILPTASTSASVSTOVEPEPEADQQLQYQOQPSISVLSDGTVDSQG 660
Db 601 TGIKQPSASTGILPTASTSASVSTOVEPEPEADQQLQYQOQPSISVLSDGTVDSQG 660
Oy 661 SSVFTESRVSSQOTVSGSQHEQASHSTGTVPBGHLPSTVQAQSQPHGYVPSSVAQSQSQG 720
Db 661 SSVFTESRVSSQOTVSGSQHEQASHSTGTVPBGHLPSTVQAQSQPHGYVPSSV----- 713
Oy 721 QPSSSILTVSSSQPIQHQQOQGIQOTAPRQOVVYSLQSTSSSEATTQAPVQAP 780
Db 721 QPSSSILTVSSSQPIQHQQOQGIQOTAPRQOVVYSLQSTSSSEATTQAPVQAP 780
Oy 781 QVLPQVSAKQ-----QGGIQOTAPRQOVVYSLQSTSSSEATTQAPVQAP 752
Db 781 QVLPQVSAKQ-----QGGIQOTAPRQOVVYSLQSTSSSEATTQAPVQAP 752
Oy 753 QVLPQVSAKQPLVPSQVPTIGRBPQIPVATQPSVVPVHSAHPLPVGQPLPTLLPOXP 812
Db 753 QVLPQVSAKQPLVPSQVPTIGRBPQIPVATQPSVVPVHSAHPLPVGQPLPTLLPOXP 812
Oy 792 ----- 791
Db 792 ----- 791
Oy 813 VSQIPISIPHVSTAQTGSSSLPITMAAGITQPLTLTSSATTAAIPGVSTVPSQLPTLL 872
Db 813 VSQIPISIPHVSTAQTGSSSLPITMAAGITQPLTLTSSATTAAIPGVSTVPSQLPTLL 872
Oy 792 ----- 791
Db 873 QPVTQPSQVHPLLQPAVQSMGIPALMGAAEVLSSGDVLVYQGFPRLLPQYIPGDSNI 932
Oy 792 ----- 791
Db 933 APSSNVASVCTHSTVLAIPMPTEVLATPGVPTVQPVVESNLLVPMGVGVGQVQVQPG 992
Oy 792 -----STQGVQVAPAEPAVAVQAPQATQPTTLTSSVDSASHVDSASMS 834
Db 993 GSILAQPTTSSQOAVLESTQGVSVAPAEPAVAVQAPQATQPTTLTSSVDSASHVDSASMS 1052
Oy 835 DGENNVPSSSGRHGRRTTKRHYRSVSRSHKTSRPKLIILVNSNGDAVVEGQLETH 894
Db 1053 DGENNVPSSSGRHGRRTTKRHYRSVSRSHKTSRPKLIILVNSNGDAVVEGQLETH 1112
Oy 895 NRKAVTFKFDLDGDNPEEIIATVNNDFIILAIERESFVQVREIIEKADMLSEBVSEP 954
Db 1113 NRKAVTFKFDLDGDNPEEIIATVNNDFIILAIERESFVQVREIIEKADMLSEBVSEP 1172
Oy 955 EGDGLESLOQKDDYGFSGSQKLGEFRKQPIPASMPQOIGIPSSSLTQVHSAKGRRTIV 1014

QY 361 TGSVYIGDLGLATLKRASFPAKSVIGTPEFMAPEMYEKKYDESVDVYAFGMCMLEMAISEY 420
 Db 301 TGSVYIGDLGLATLKRASFPAKSVIGTPEFMAPEMYEKKYDESVDVYAFGMCMLEMAISEY 360
 QY 421 PYSECQNAAQIYRRYTSVGPASFDKVAIPVKEIIEGCIQNOKDERYSIKDILNHAFFQ 480
 Db 361 PYSECQNAAQIYRRYTSVGPASFDKVAIPVKEIIEGCIQNOKDERYSIKDILNHAFFQ 420
 QY 481 EETGVARELAEDDEGKAIKILMLRIEDIKLKGKYKONNEAIEFSPDERVPEDEVQEM 540
 Db 421 EETGVARELAEDDEGKAIKILMLRIEDIKLKGKYKONNEAIEFSPDERVPEDEVQEM 480
 QY 541 VESGYVEGDHKTAKAKIKORVSLIKRREORQLVREOEKKKOEBSLKOQVQSSASQ 600
 Db 481 VESGYVEGDHKTAKAKIKORVSLIKRREORQLVREOEKKKOEBSLKOQVQSSASQ 540
 QY 601 TGKIQLPASASTGTTASTTASVSTQVPEPEADQHQLOQQPSISVLSGTVDSGQG 660
 Db 541 TGKIQLPASASTGTTASTTASVSTQVPEPEADQHQLOQQPSISVLSGTVDSGQG 600
 QY 661 SSVFTESSVSOQTVSQSGOHEOAHSTGVPGHIPSITQAOQOPHGYVPPSSVAQOGSOG 720
 Db 601 SSVFTESSVSOQTVSQSGOHEOAHSTGVPGHIPSITQAOQOPHGYVPPSSVAQOGSOG 660
 QY 721 QPSSSLTGVSSSQPIQHPQQOQGIQQTAPPOQTVOYSLSQTSSEATTAQPPVSOQAP 780
 Db 661 QPSSSLTGVSSSQPIQHPQQOQGIQQTAPPOQTVOYSLSQTSSEATTAQPPVSOQAP 720
 QY 781 QVLPQVSAKOSTQGVSOQVAPAEVAVAAQOPATQPTTLASVDSASDVASGMSDGENV 840
 Db 721 QVLPQVSAKOSTQGVSOQVAPAEVAVAAQOPATQPTTLASVDSASDVASGMSDGENV 780
 QY 841 PSSSGRHEGRRTTKHYKRSVSRHEKTSRPKRLINLSNKGRRVYECQLETNRKMYT 900
 Db 781 PSSSGRHEGRRTTKHYKRSVSRHEKTSRPKRLINLSNKGRRVYECQLETNRKMYT 840
 QY 901 FKFDLDGDNPEEIAITIMWNDPILAIERESFVDOVREIIEKADMLSEDVSVEEBQGL 960
 Db 841 FKFDLDGDNPEEIAITIMWNDPILAIERESFVDOVREIIEKADMLSEDVSVEEBQGL 900
 QY 961 ESLOGKDDYGGSSQKLEGEFKQPIPASSMPQOIGIPTSSITQVYVHSAGRFFVSPVES 1020
 Db 901 ESLOGKDDYGGSSQKLEGEFKQPIPASSMPQOIGIPTSSITQVYVHSAGRFFVSPVES 960
 QY 1021 RLRSKVPFSEITQTVASTASQPMNLSSHASSLSIQAFSELRRAQMTGEPATAPNF 1080
 Db 961 RLRSKVPFSEITQTVASTASQPMNLSSHASSLSIQAFSELRRAQMTGEPATAPNF 1020
 QY 1081 SHTGPTFPVVVPPFLSSIAAGVPTTAATAPVATSSPPNDISTSVIOSEVTVTEEGIAGV 1140
 Db 1021 SHTGPTFPVVVPPFLSSIAAGVPTTAATAPVATSSPPNDISTSVIOSEVTVTEEGIAGV 1080
 QY 1141 ATSTGVVTSGGGLPIPPVSESPVLSVVSITTPAVVSIISTTSPSLQVPTSEIIVSSTA 1200
 Db 1081 ATSTGVVTSGGGLPIPPVSESPVLSVVSITTPAVVSIISTTSPSLQVPTSEIIVSSTA 1140
 QY 1201 LYPSTVATSASAGSSTATPGPKPPAVVSOQAAGSTTVGATILSVSTTSSFPETAQOLS 1260
 Db 1141 LYPSTVATSASAGSSTATPGPKPPAVVSOQAAGSTTVGATILSVSTTSSFPETAQOLS 1200
 QY 1261 IQSSSSTSTPTLATVTVVSAHSLDKTSHSSTTGLAFSLASAPSSSSSGAGVSSISIQPGG 1320
 Db 1201 IQSSSSTSTPTLATVTVVSAHSLDKTSHSSTTGLAFSLASAPSSSSSGAGVSSISIQPGG 1260
 QY 1321 LHPVIVSVIASVPIPLFOAGPTSTPLIPQVPSIPPLVQPVANYPANVQOQTLIHSQOPAL 1380
 Db 1261 LHPVIVSVIASVPIPLFOAGPTSTPLIPQVPSIPPLVQPVANYPANVQOQTLIHSQOPAL 1320
 QY 1381 LPNQPHTHCEVDDTQPKAPGIDDKITLSEKLSISESHSSSGAQAHSVLSFSTVYES 1440
 Db 1321 LPNQPHTHCEVDDTQPKAPGIDDKITLSEKLSISESHSSSGAQAHSVLSFSTVYES 1380

QY 1441 TVTPGIEPTTAVAPSKLITSTTSTCLPPTNLPLGVALPVTVPGQVSTPVSTTTSQVK 1500
 Db 1381 TVTPGIEPTTAVAPSKLITSTTSTCLPPTNLPLGVALPVTVPGQVSTPVSTTTSQVK 1440
 QY 1501 PGTAPSKPEPLTKAVPLPVGTIELPAGTILPSEQLPPFPQPSLTSQOQPLEDDLAQLRRTLSF 1560
 Db 1441 PGTAPSKPEPLTKAVPLPVGTIELPAGTILPSEQLPPFPQPSLTSQOQPLEDDLAQLRRTLSF 1500
 QY 1561 EXITVTSVAVPVSAAATATTEAGTOPOKVSOVKEBPVLTATSSGACVFKKGRQVSVAA 1620
 Db 1501 EXITVTSVAVPVSAAATATTEAGTOPOKVSOVKEBPVLTATSSGACVFKKGRQVSVAA 1560
 QY 1621 DGAQKEGKNKSEDAKSVHFESESTSESSVLSSSPSESTLVKPEPGITIPGISDVPESSAH 1680
 Db 1561 DGAQKEGKNKSEDAKSVHFESESTSESSVLSSSPSESTLVKPEPGITIPGISDVPESSAH 1620
 QY 1681 KTTASEAKSDTQPTKYGRFOVTTTANKVGRFSVSKTEDKITDTYKKGVPVAPPPMDLEQ 1740
 Db 1621 KTTASEAKSDTQPTKYGRFOVTTTANKVGRFSVSKTEDKITDTYKKGVPVAPPPMDLEQ 1680
 QY 1741 AVLPAAVLPKKEKPELSPSHLNGPSSDPEAAFLSRDVEDGSGSPHOLSSKSLPSONL 1800
 Db 1681 AVLPAAVLPKKEKPELSPSHLNGPSSDPEAAFLSRDVEDGSGSPHOLSSKSLPSONL 1740
 QY 1801 SOSLSNSFNSSVMSDNESEDIEDEDLKLIELRLRDKHLKEIQDQOSRQKHEISLYTKLG 1860
 Db 1741 SOSLSNSFNSSVMSDNESEDIEDEDLKLIELRLRDKHLKEIQDQOSRQKHEISLYTKLG 1800
 QY 1861 KVPAAVITPPAALPSGRRRPTYSKSGSSRSLSLGNKSPQLSGNLSGQSAASVLAHQOT 1920
 Db 1801 KVPAAVITPPAALPSGRRRPTYSKSGSSRSLSLGNKSPQLSGNLSGQSAASVLAHQOT 1860
 QY 1921 LHPGNTPESGONOLQPLKPSPSDNLVYAFISDGAISVPSISAPQOGTSSITTGATV 1980
 Db 1861 LHPGNTPESGONOLQPLKPSPSDNLVYAFISDGAISVPSISAPQOGTSSITTGATV 1920
 QY 1981 NSQAQAQAPPAAMTSSRKGTFTDDLHKLVDNWARDAMNLSGRGSGKGMNVEGPMARKFS 2040
 Db 1921 NSQAQAQAPPAAMTSSRKGTFTDDLHKLVDNWARDAMNLSGRGSGKGMNVEGPMARKFS 1980
 QY 2041 APGOLCISMTSNUGSAPISAASATSLGHTTKSMCPQOQIGFPATPPGAOMSGTGAPAQ 2100
 Db 1981 APGOLCISMTSNUGSAPISAASATSLGHTTKSMCPQOQIGFPATPPGAOMSGTGAPAQ 2040
 QY 2101 PLGQFQPVGTASLONENISNLQKSIINPPGSNLRTT 2136
 Db 2041 PLGQFQPVGTASLONENISNLQKSIINPPGSNLRTT 2100
 RESULT 11
 ID ABO44400 standard; protein: 2169 AA.
 XX ABO44400;
 AC ABO44400;
 DT 26-SEP-2003 (first entry)
 DE Novel human protein kinase #20.
 KW Human; kinase; enzyme; cosmetic application; nutraceutical application.
 OS Homo sapiens.
 XX US6541252-B1.
 XX 01-APR-2003.
 XX PD 14-MAY-2001; 2001US-00854856.
 XX PF 19-MAY-2000; 2000US-0206015P.
 XX PR
 XX PA (LEXI-) LEXICON GENETICS INC.
 XX

PI Walke DW, Hilbun E, Donoho G, Turner CA;

XX MPI: 2003-575927/54.

DR N-PSDB; ACH03783.

XX New nucleic acid encoding novel human proteins, useful in cosmetic and
PT nutritional applications.

XX Disclosure; Page; 11pp; English.

CC The invention relates to a new isolated nucleic acid encoding a novel
CC human protein kinase. The nucleic acid is useful in cosmetic and
CC nutritional applications. The present sequence represents the amino
CC acid sequence of a novel human protein kinase. Note: The sequence data
CC for this patent did not form part of the printed specification but was
CC obtained in electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?docid=6541252B1

XX Sequence 2169 AA;

Query Match 96.7%; Score 10451.5; DB 7; Length 2169;

Best local similarity 95.7%; Pred. No. 0; Indels 93; Gaps 1;

Matches 2076; Conservative 0; Mismatches 0; Indels 93; Gaps 1;

QY 61 MDKSRGAATTTTTEHFRFRSVICDSNATALELPLSLPQPSIPAAVPOSAPPPEH 120
DB 1 MDKSRGAATTTTTEHFRFRSVICDSNATALELPLSLPQPSIPAAVPOSAPPPEH 60
QY 121 REETVATATISQVAOQPPAAAPGQAVAGPAPSTVPSTSKDRPVSGPSLVGSKKEPP 180
DB 61 REETVATATISQVAOQPPAAAPGQAVAGPAPSTVPSTSKDRPVSGPSLVGSKKEPP 120
QY 181 ARSGSGGSAKEPEERSQODDIELEETKAVGMSNDKFLKPIELGRGSEKTVYGLD 240
DB 121 ARSGSGGSAKEPEERSQODDIELEETKAVGMSNDKFLKPIELGRGSEKTVYGLD 180
QY 241 TETVEVAMCELQDRKLTSEKORFKEEAEMKGLQHPNIVRFYDSMESTVKGKCIYLV 300
DB 181 TETVEVAMCELQDRKLTSEKORFKEEAEMKGLQHPNIVRFYDSMESTVKGKCIYLV 240
QY 301 TELMTSGTLKTYLQKFKVMKIKVLRSMCRQILKGLQFLHTTPPIIHRDLKCDNIFITGP 360
DB 241 TELMTSGTLKTYLQKFKVMKIKVLRSMCRQILKGLQFLHTTPPIIHRDLKCDNIFITGP 300
QY 361 TGSVKIGLGLATIKRASFASVIGTPEFMAPEMVEEKYDESVDVYAFGMCMLEMTSEY 420
DB 301 TGSVKIGLGLATIKRASFASVIGTPEFMAPEMVEEKYDESVDVYAFGMCMLEMTSEY 360
QY 421 PYSECQNAOQYRRVTSQVGFVDFKVAIPVKEIIEGCIKONDERYSIKDLINHAFFQ 480
DB 361 PYSECQNAOQYRRVTSQVGFVDFKVAIPVKEIIEGCIKONDERYSIKDLINHAFFQ 420
QY 481 BETGVAVELAEEDDEGEKIAIKLMLRIEDIKLKGKYNENAIIEFSFDLERDVEDVAQEM 540
DB 421 BETGVAVELAEEDDEGEKIAIKLMLRIEDIKLKGKYNENAIIEFSFDLERDVEDVAQEM 480
QY 541 VESGVYCEGDHMKAKAIKORVSLIKRRRQRLVRESEKKKQESLTKOVVSSASQ 600
DB 481 VESGVYCEGDHMKAKAIKORVSLIKRRRQRLVRESEKKKQESLTKOVVSSASQ 540
QY 601 TGIKOLPASASTGISTASVSTOPEPEPEADQHQLOQYQPSISVLSDGTVDGSGG 660
DB 541 TGIKOLPASASTGISTASVSTOPEPEPEADQHQLOQYQPSISVLSDGTVDGSGG 600
QY 661 SSVTETESRVSQOQVVSQSHQAHSTGTUGHLPSTVYQASQPHGYPPSSVAQSGSG 720
DB 601 SSVTETESRVSQOQVVSQSHQAHSTGTUGHLPSTVYQASQPHGYPPSSVAQSGSG 660
QY 721 QPSSSSTLVGSSSOPIDHPQOQOGLIOTAPPOQTVOYSLQTSSTRAATTAQVPSQAP 780
DB 661 QPSSSSTLVGSSSOPIDHPQOQOGLIOTAPPOQTVOYSLQTSSTRAATTAQVPSQAP 720
QY 781 QVLPQVSAKQO----- 791

DB 721 QVLPQVSAKQOFPPLPPOYPGDSNIAPSNNAVCIIHSTVLKXPMFTEVLATPGREPT 780
QY 792 -----STGVSGVAPAPVAV 807
DB 781 VVQPYVESNLLVPMGVGGVQVQSGPGSLAQAPLTSSQQAVALSTQGVSGVAPAPVAV 840
QY 808 AQPOATQPTTLASVDASHVAGMDGNENNVSSSGRHEGRITTKHRYKSVSRSHHE 867
DB 841 AQPOATQPTTLASVDASHVAGMDGNENNVSSSGRHEGRITTKHRYKSVSRSHHE 808
QY 868 KTSAPKRIINLVNKKGRVVECCLETHNRKMTVTFKFDLQGNPEIATIMNNPILIE 927
DB 901 KTSAPKRIINLVNKKGRVVECCLETHNRKMTVTFKFDLQGNPEIATIMNNPILIE 868
QY 928 RESFVDQREIIEKADMLSEDSVSEPEGDGLSLQKDDYGFSGSKLEGEFKQPIPA 987
DB 961 RESFVDQREIIEKADMLSEDSVSEPEGDGLSLQKDDYGFSGSKLEGEFKQPIPA 928
QY 988 SSMPOQIGIPSSLTQVHSAGRFIVSPVESRLRSKVPSEITDVAASTAQSPMN 1047
DB 1021 SSMPOQIGIPSSLTQVHSAGRFIVSPVESRLRSKVPSEITDVAASTAQSPMN 988
QY 1048 LSHSASSLSIQOARSELRAQMTGEPNTAPENFSGHTGTFPVVPPFLSSIAGVPTTAAT 1107
DB 1081 LSHSASSLSIQOARSELRAQMTGEPNTAPENFSGHTGTFPVVPPFLSSIAGVPTTAAT 1048
QY 1108 APVPATSSPNDISTSVIOSEVITPTEEGIAAGVATSGVTSQGLPIPVSESVLSSV 1167
DB 1141 APVPATSSPNDISTSVIOSEVITPTEEGIAAGVATSGVTSQGLPIPVSESVLSSV 1108
QY 1168 SSTIPAVASISTSPLOVPTSEIIVSSSTALYPVATSAASAGSATGPKRPA 1227
DB 1201 SSTIPAVASISTSPLOVPTSEIIVSSSTALYPVATSAASAGSATGPKRPA 1168
QY 1228 VVSGQAAGSTTVGATLTVSTSTTSFPESTASQSLQLSSTSTPTLAETVVAASHLDTKS 1287
DB 1261 VVSGQAAGSTTVGATLTVSTSTTSFPESTASQSLQLSSTSTPTLAETVVAASHLDTKS 1228
QY 1288 HSTTTGIAFCSIAPESSSSSGAGVSVISQPGCHLPVIBSVIATSTPLPAAQPTSTPL 1347
DB 1321 HSTTTGIAFCSIAPESSSSSGAGVSVISQPGCHLPVIBSVIATSTPLPAAQPTSTPL 1288
QY 1348 LPOVPSIPPLVOPANVPAVOQTLIHQOPALPNOPHHCPEVDSTQKAGIDIK 1407
DB 1381 LPOVPSIPPLVOPANVPAVOQTLIHQOPALPNOPHHCPEVDSTQKAGIDIK 1348
QY 1408 TLEBKLSLSEHSSSGAASHVSLSTSVLESTVTEGIPPTTAAVAPSKLITSTSTCLP 1467
DB 1441 TLEBKLSLSEHSSSGAASHVSLSTSVLESTVTEGIPPTTAAVAPSKLITSTSTCLP 1408
QY 1468 TNLPLGTVALPVTPVTPGQVSTVSTTTGKVKGTAPSKRPLTKAPVLPYGTBLPAGTL 1527
DB 1501 TNLPLGTVALPVTPVTPGQVSTVSTTTGKVKGTAPSKRPLTKAPVLPYGTBLPAGTL 1468
QY 1528 PSBQLPFPGLSLQSQOPLLEDLADQLRFLTSPKXITVTSVAGVSNMAAPATREAGTQ 1587
DB 1561 PSBQLPFPGLSLQSQOPLLEDLADQLRFLTSPKXITVTSVAGVSNMAAPATREAGTQ 1528
QY 1588 QKGVSYKEGFLVATSSGAGVFKMGRPOVVAADQAKGKNKSEDAKSVHESSTSESS 1647
DB 1621 QKGVSYKEGFLVATSSGAGVFKMGRPOVVAADQAKGKNKSEDAKSVHESSTSESS 1588
QY 1648 VLSSSSPESTLVKRENGITIPGISLSVPESAAKTTTSAKSDTQGTKYGRPOVTTAN 1707
DB 1681 VLSSSSPESTLVKRENGITIPGISLSVPESAAKTTTSAKSDTQGTKYGRPOVTTAN 1648
QY 1708 KVGFSYSKTEDEKTTDRKKGPPVAPPFMDLEQAVLPAVLPKKEKPELSEPSHLNGSSSD 1767
DB 1741 KVGFSYSKTEDEKTTDRKKGPPVAPPFMDLEQAVLPAVLPKKEKPELSEPSHLNGSSSD 1708
QY 1768 PEAFLSRDVEDGSGSPHSQOLSKSLPSQNTLSQSLNENSSYMSDNESEDIEDLK 1827

Db 1801 PEAALFSDVDGSGSPHSPOHLSKSLPQNLGSLNSFNSSYMSSDNEEDLK 1860
 QY 1828 LELRLRLRDLKLEIIDLOROKHEIESLYTLGKVPVAVIIPPAAPLSGRRRPTKSGS 1887
 Db 1861 LELRLRLRDLKLEIIDLOROKHEIESLYTLGKVPVAVIIPPAAPLSGRRRPTKSGS 1920
 QY 1888 KSSRSSSLGKNSPOLSGNLGSGOASAVLHPQOTLHPNIPESGONLLQELKPSPSDN 1947
 Db 1921 KSSRSSSLGKNSPOLSGNLGSGOASAVLHPQOTLHPNIPESGONLLQELKPSPSDN 1980
 QY 1948 LYSATFSTGALSIVSLSPAGCGTSTNTVGATVNSQAQAOPPMATSRKCTFTDDLHL 2007
 Db 1981 LYSATFSDGALSIVSLSPAGCGTSTNTVGATVNSQAQAOPPMATSRKCTFTDDLHL 2040
 QY 2008 VDNARADAMNLGSRGSKGHMNVGPGWAKKFSAPGOLCISMTSNLGGADISAASATSL 2067
 Db 2041 VDNARADAMNLGSRGSKGHMNVGPGWAKKFSAPGOLCISMTSNLGGADISAASATSL 2100
 QY 2068 GHFTKSMCPPOQYGFPAIPFGAOWSGTGPAPOPPLQGPQVPVGTASLQNFNISLQKSIISN 2127
 Db 2101 GHFTKSMCPPOQYGFPAIPFGAOWSGTGPAPOPPLQGPQVPVGTASLQNFNISLQKSIISN 2160
 QY 2128 PPGSNLRTT 2136
 Db 2161 PPGSNLRTT 2169
 RESULT 12
 ABO44397
 ID ABO44397 standard; protein; 2322 AA.
 XX ABO44397;
 AC 26-SEP-2003 (first entry)
 XX 26-SEP-2003 (first entry)
 DE Novel human protein kinase #17.
 XX Human; kinase; enzyme; cosmetic application; nutraceutical application.
 KW Homo sapiens.
 OS US6541252-B1.
 EN 01-APR-2003.
 PD 14-MAY-2001; 2001US-00854856.
 PF 19-MAY-2000; 2000US-0206015P.
 ER (LEXI-) LEXICON GENETICS INC.
 XX PA
 PI Walke DW, Hilbun E, Donoho G, Turner CA;
 XX WPI; 2003-575927/54.
 DR N-PSDB; ACH03780.
 XX
 PT New nucleic acid encoding novel human proteins, useful in cosmetic and
 XX nutraceutical applications.
 PS Disclosure; Page; 11pp; English.
 CC The invention relates to a new isolated nucleic acid encoding a novel
 CC human protein kinase. The nucleic acid is useful in cosmetic and
 CC nutraceutical applications. The present sequence represents the amino
 CC acid sequence of a novel human protein kinase. Note: The sequence data
 CC for this patent did not form part of the printed specification but was
 CC obtained in electronic format directly from USPTO at
 CC seqdata.uspto.gov/sequence.html?docID=6541252B1
 XX
 SQ Sequence 2322 AA:
 Query Match 96.0%; Score 10375; DB 7; Length 2322;
 Best Local Similarity 89.4%; Pred. No. 0;

Matches 2076; Conservative 0; Mismatches 0; Indels 246; Gaps 1;
 QY 61 MDKSRGAAATTTTTERFRFRSVICDNNATLLEPLGLPLSLPQPSIPAAVPOGAPPEPH 120
 Db 1 MDKSRGAAATTTTTERFRFRSVICDNNATLLEPLGLPLSLPQPSIPAAVPOGAPPEPH 60
 QY 121 REETVATATSOVAQAPPAAPAAPEQAVAGAPSTVSTSKDBPVSOPLVSGKEEPP 180
 Db 61 REETVATATSOVAQAPPAAPAAPEQAVAGAPSTVSTSKDBPVSOPLVSGKEEPP 120
 QY 181 ARSSGGGSAKEPEEESQOQDDIELEETAVAGSNGRFLKFTIEIGRGFKTVYKGLD 240
 Db 121 ARSSGGGSAKEPEEESQOQDDIELEETAVAGSNGRFLKFTIEIGRGFKTVYKGLD 180
 QY 241 TETVEVAMCELOQRKLTTSRPFKEAEMLKGLQHPNIYRFVDSNESTYKGGKCIYLV 300
 Db 181 TETVEVAMCELOQRKLTTSRPFKEAEMLKGLQHPNIYRFVDSNESTYKGGKCIYLV 240
 QY 301 TELMTSGTLKTYLKRFRVMKIKVLRSMCROILKGLQELHTPTPIIHRDLKCDNIPTTGP 360
 Db 241 TELMTSGTLKTYLKRFRVMKIKVLRSMCROILKGLQELHTPTPIIHRDLKCDNIPTTGP 300
 QY 361 TGSVKIGDLGIATIKRASPAKSVIGTEPFMAPEVTEKYEVSUVVYAFGCMLEMAVSEY 420
 Db 301 TGSVKIGDLGIATIKRASPAKSVIGTEPFMAPEVTEKYEVSUVVYAFGCMLEMAVSEY 360
 QY 421 PYSECQNAAOIYRVTSGVXPASFQKVAIPVKEIIEGCIKRONDERYSIDLINHAFFQ 480
 Db 361 PYSECQNAAOIYRVTSGVXPASFQKVAIPVKEIIEGCIKRONDERYSIDLINHAFFQ 420
 QY 481 EETGVAVELAEEDDGEKIAIKLWLRIDIKLKQKYDNEAIEFSFLERDVPEDVAQEM 540
 Db 421 EETGVAVELAEEDDGEKIAIKLWLRIDIKLKQKYDNEAIEFSFLERDVPEDVAQEM 480
 QY 541 VESGVCEGDHKTMAKAIKDRVSLIKRREROLVREOEKKOEBSLKOQVQSSASQ 600
 Db 481 VESGVCEGDHKTMAKAIKDRVSLIKRREROLVREOEKKOEBSLKOQVQSSASQ 540
 QY 601 TGIKQLPSASTGIPYATTSASVSTQVEPEEPADQHQLOLOQOPSTISVLSGTVDSGQ 660
 Db 541 TGIKQLPSASTGIPYATTSASVSTQVEPEEPADQHQLOLOQOPSTISVLSGTVDSGQ 600
 QY 661 SSVFTESRVSSQOTVSYGSGHEQASHSTGVPGHIPSTVQAQSGHGVPPSSVAQOGSQ 720
 Db 601 SSVFTESRVSSQOTVSYGSGHEQASHSTGVPGHIPSTVQAQSGHGVPPSSVAQOGSQ 660
 QY 721 QPSSSILTVSSSQPIQHPOQOQCIQOTAPQOQVQVSLQSTSTSEATTNQPVSQOPAP 780
 Db 661 QPSSSILTVSSSQPIQHPOQOQCIQOTAPQOQVQVSLQSTSTSEATTNQPVSQOPAP 720
 QY 781 QVLPOVSAGKQ----- 791
 Db 721 QVLPOVSAGKQLPVSPVPTTQGEPOIPVATQPSVVPVHSGAHLPVGQPLPTLLPOXP 780
 QY 791 ----- 791
 Db 781 VSGIPISTPHVSTAQTGSSLPITMAAGITQPLITLASSATTAAIPGVSTVPESQLPTLL 840
 QY 792 ----- 791
 Db 841 QPVTQPSQVHPOLLQPAVQSGMGIPIANLQAAEVLSSGDLVLYGFPPLRPQYRGDSNI 900
 QY 792 ----- 791
 Db 901 APSSNVASVCTHSTVLRPMTVELATPGFPTVQVPVESNLVLPMGVGQVQVQSPG 960
 QY 792 -----STGVQSVAAPEAPVAVAAQPOATFTTTLASSVDSASHSDVASGMS 834
 Db 961 GSLAQPPTSSQAAVLESTQGVSVQVAPAEVAVAAQPOATFTTTLASSVDSASHSDVASGMS 1020
 QY 835 DGNENVPSSGRRHGRTTKKHYRKSVPSSRSHKTSRKLILNVSNKGDVAVVCOLETH 894
 Db 1021 DGNENVPSSGRRHGRTTKKHYRKSVPSSRSHKTSRKLILNVSNKGDVAVVCOLETH 1080

```

QY 895 NRKXWTFKFDLDGDNPERIATIMVNDPFLIAERESFVDQVRELLKXADEMLSENVSEP 954
Db 1081 NRKXWTFKFDLDGDNPERIATIMVNDPFLIAERESFVDQVRELLKXADEMLSENVSEP 1140
QY 955 EGDQGLJSLQCKNDYGFSGSQCLBGEFKQPIPASSMPOQIGIPISLTQVHSAGRREIV 1014
Db 1141 EGDQGLJSLQCKNDYGFSGSQCLBGEFKQPIPASSMPOQIGIPISLTQVHSAGRREIV 1200
QY 1015 SPVPSRLRESKVFSEITDVTAASTAOSPGMNLSHSASLSLQOAFSELRAQMTBGN 1074
Db 1201 SPVPSRLRESKVFSEITDVTAASTAOSPGMNLSHSASLSLQOAFSELRAQMTBGN 1260
QY 1075 TAPPNFSHTGTFEPVPPFLSIAGVPTTAATAATAPVATSSPNDISTSVIOSEVTEPTE 1134
Db 1261 TAPPNFSHTGTFEPVPPFLSIAGVPTTAATAATAPVATSSPNDISTSVIOSEVTEPTE 1320
QY 1135 EGIAGVATSTGVTSGGLPIPVSESPLYSSVSSITTPAVVISITTSPSIQVPTSTSEI 1194
Db 1321 EGIAGVATSTGVTSGGLPIPVSESPLYSSVSSITTPAVVISITTSPSIQVPTSTSEI 1380
QY 1195 VVSSSTALPSTVTSATASAGSSTATPGKPPAVVSOQAAGSTTVGATLTSVSTTSFPS 1254
Db 1381 VVSSSTALPSTVTSATASAGSSTATPGKPPAVVSOQAAGSTTVGATLTSVSTTSFPS 1440
QY 1255 TASQLSIQSSSTSTPTLAETVWVAHSLDKTSHSTTGALFSLAPSSSSPGAGVSSY 1314
Db 1441 TASQLSIQSSSTSTPTLAETVWVAHSLDKTSHSTTGALFSLAPSSSSPGAGVSSY 1500
QY 1315 ISOPGGLHPLVIPSVAIATPILPQAAGPTSTPLPQVPSIPPLVQPVANVAVOQTLIHS 1374
Db 1501 ISOPGGLHPLVIPSVAIATPILPQAAGPTSTPLPQVPSIPPLVQPVANVAVOQTLIHS 1560
QY 1375 QOPRALPNOGHTHCPEDVSDTOPKAPGIDIKITEELKRLFSHSSSGAQHSAVSLET 1434
Db 1561 QOPRALPNOGHTHCPEDVSDTOPKAPGIDIKITEELKRLFSHSSSGAQHSAVSLET 1620
QY 1435 SLVLESTVTGIPITTAVAVSKLITSTSTCLPNTLPGTVALPTPVTEGGVSTPST 1494
Db 1621 SLVLESTVTGIPITTAVAVSKLITSTSTCLPNTLPGTVALPTPVTEGGVSTPST 1680
QY 1495 TTSGVKEGTABSKPPLTKAPVLPVGTLPSEQLPFPGPSLITGSOQPLEDLAQ 1554
Db 1681 TTSGVKEGTABSKPPLTKAPVLPVGTLPSEQLPFPGPSLITGSOQPLEDLAQ 1740
QY 1555 RRTISPEXITVTSAGVPVMAAPTAITFAGTQPKQVSGVGEVYLATSSGAGVFKMGFR 1614
Db 1741 RRTISPEXITVTSAGVPVMAAPTAITFAGTQPKQVSGVGEVYLATSSGAGVFKMGFR 1800
QY 1615 QVSVAAADGAQKEGKNKSEDAKSVHFESTSESSVLSSSSPSTLVKPEPNCITTPGISD 1674
Db 1801 QVSVAAADGAQKEGKNKSEDAKSVHFESTSESSVLSSSSPSTLVKPEPNCITTPGISD 1860
QY 1675 VPESAHKTITASAEASDTCQPTKVPQVYTTANKVGRFSVSKTEDKITDTKKGEPVAP 1734
Db 1861 VPESAHKTITASAEASDTCQPTKVPQVYTTANKVGRFSVSKTEDKITDTKKGEPVAP 1920
QY 1735 FMDLEQAVLPVITPKKEPELSEPSHLNGSPDPEAAFLSDVDDGSSPHSPHQLSSKS 1794
Db 1921 FMDLEQAVLPVITPKKEPELSEPSHLNGSPDPEAAFLSDVDDGSSPHSPHQLSSKS 1980
QY 1795 LPSQNLQSLSNSFNSSSDNESDIEDLKLRLRDKHKEIIOLOSROKHETES 1854
Db 1981 LPSQNLQSLSNSFNSSSDNESDIEDLKLRLRDKHKEIIOLOSROKHETES 2040
QY 1855 LYTYLGVKVPVAVIIPPAAPLSGRRRPTKSGKSSRSSSLGNKSPOLSGNLSQCSASV 1914
Db 2041 LYTYLGVKVPVAVIIPPAAPLSGRRRPTKSGKSSRSSSLGNKSPOLSGNLSQCSASV 2100
QY 1915 LHPQOOLHPGONTPESSONOLLOPLKSPSSDNTYSAFTSGAISVPSLSAPGCGTSTN 1974
Db 2101 LHPQOOLHPGONTPESSONOLLOPLKSPSSDNTYSAFTSGAISVPSLSAPGCGTSTN 2160

```

```

QY 1975 TVGATVNSQAQAOPPAMTSSRKGTFTDDLHKLVDNWARDAMNLSGRGSKHMYEGPG 2034
Db 2161 TVGATVNSQAQAOPPAMTSSRKGTFTDDLHKLVDNWARDAMNLSGRGSKHMYEGPG 2220
QY 2035 MARKFSAPGOLCISMTSNLGGSAPIASAATSLSLGHFTSMCPPOOYGFPAIPFGAOMSGT 2094
Db 2221 MARKFSAPGOLCISMTSNLGGSAPIASAATSLSLGHFTSMCPPOOYGFPAIPFGAOMSGT 2280
QY 2095 GGPAPOPUGOPQVGTASLQNFNINSLQKSIISNPPGSNLRTT 2136
Db 2281 GGPAPOPUGOPQVGTASLQNFNINSLQKSIISNPPGSNLRTT 2322

RESULT 13
ABO44411
ID ABO44411 standard; protein; 2048 AA.
XX
XX ABO44411;
AC
XX
XX 26-SEP-2003 (first entry)
DT
XX
XX Novel human protein kinase #31.
DE
XX
XX Human; kinase; enzyme; cosmetic application; nutraceutical application.
KW
XX
XX Homo sapiens.
OS
XX
XX US6541252-B1.
EN
XX
XX 01-APR-2003.
PD
XX
XX 14-MAY-2001; 2001US-00854856.
PR
XX
XX 19-MAY-2000; 2000US-0206015P.
PR
XX
XX (LEXI-) LEXICON GENETICS INC.
PA
XX
XX
PI
XX
XX Walke DW, Hilbun E, Donoho G, Turner CA;
XX
XX WPI; 2003-575927/54.
DR
XX
XX N-PSDB; ACH03794.
XX
XX
XX New nucleic acid encoding novel human proteins, useful in cosmetic and
PT nutraceutical applications.
PS
XX
XX Disclosure; Page: 11pp; English.
PS
XX
XX The invention relates to a new isolated nucleic acid encoding a novel
CC human protein kinase. The nucleic acid is useful in cosmetic and
CC nutraceutical applications. The present sequence represents the amino
CC acid sequence of a novel human protein kinase. Note: The sequence data
CC for this patent did not from part of the printed specification but was
CC obtained in electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?docid=6541252B1
XX
XX
SQ Sequence 2048 AA;

Query Match 95.7%; Score 10344; DB 7; Length 2048;
Best Local Similarity 98.7%; Pred. No. 0;
Matches 2048; Conservative 0; Mismatches 0; Indels 28; Gaps 1;
QY 61 MDKOSRGAATTTTTHRRFRFRSVTCDSNATALELPGLPLSLPQPSIPAAVPOSAPPEPH 120
Db 1 MDKOSRGAATTTTTHRRFRFRSVTCDSNATALELPGLPLSLPQPSIPAAVPOSAPPEPH 60
QY 121 REETVTATATSOVAQCPAAAPAEQAVAGPAPSTVPSSTKDRPVSOPLVSGKEPP 180
Db 61 REETVTATATSOVAQCPAAAPAEQAVAGPAPSTVPSSTKDRPVSOPLVSGKEPP 120
QY 181 ARSGSGGSAKEPQERSQOQDDIELETAAGVMSNDGRFLKEDIETGRSGFKTVYGLD 240
Db 121 ARSGSGGSAKEPQERSQOQDDIELETAAGVMSNDGRFLKEDIETGRSGFKTVYGLD 180

```

QY 241 TETTVAWACELQNRKLTSEKQRFKEEAEMLKGLQHPNIVRFYDSWESTVKGKCIIVL 300
 Db 181 TETVEAWACELQNRKLTSEKQRFKEEAEMLKGLQHPNIVRFYDSWESTVKGKCIIVL 240
 QY 301 TELMTSGTLKTYLKGPFYMKIKVLRSMCRQILKGLQFLHTTPTPIIHRDLKCDNI FTTPG 360
 Db 241 TELMTSGTLKTYLKGPFYMKIKVLRSMCRQILKGLQFLHTTPTPIIHRDLKCDNI FTTPG 300
 QY 361 TGSVKIGDLGLATLKRAFPKSVIGTPEFMAPEMYEKEKYDESUVVYAFGMCLMATEY 420
 Db 301 TGSVKIGDLGLATLKRAFPKSVIGTPEFMAPEMYEKEKYDESUVVYAFGMCLMATEY 360
 QY 421 PYSCONAAQIYRVTSGVCPASFDKVAIPEVKEIIEGCIRQNDERYSIKDLNHAFFQ 480
 Db 361 PYSCONAAQIYRVTSGVCPASFDKVAIPEVKEIIEGCIRQNDERYSIKDLNHAFFQ 420
 QY 481 EETGVRELAEDDGEKIAIKLMLRIEDIKLKGYKQNEAIEFSFDLERDPEDVAQEM 540
 Db 421 EETGVRELAEDDGEKIAIKLMLRIEDIKLKGYKQNEAIEFSFDLERDPEDVAQEM 480
 QY 541 VESGVCEGDKHTMAKAIKDRVSLIKRREOROLVREOEKKQKQESSLKQVQSSASQ 600
 Db 481 VESGVCEGDKHTMAKAIKDRVSLIKRREOROLVREOEKKQKQESSLKQVQSSASQ 540
 QY 601 TGIRKQPSASTGIPASTTSASVSTOVEPEEPADQHQOIQYQOPISISVSDGTVDSGQ 660
 Db 541 TGIRKQPSASTGIPASTTSASVSTOVEPEEPADQHQOIQYQOPISISVSDGTVDSGQ 600
 QY 661 SSVTETSRVSSQOIVSYSGSOHEQAHSTGTVFGHLPSTYQAOOSQPHGYIPBSVAQOSQ 720
 Db 601 SSVTETSRVSSQOIVSYSGSOHEQAHSTGTVFGHLPSTYQAOOSQPHGYIPBSVAQOSQ 660
 QY 721 QPSSSILTVSSSQPIQHPOQOQGIQOTAPPOQTVOYSLSTQTSSEATTQAPVOPAP 780
 Db 661 QPSSSILTVSSSQPIQHPOQOQGIQOTAPPOQTVOYSLSTQTSSEATTQAPVOPAP 720
 QY 781 QVLPQVASAGKOSTQVSVQVAPAEFVAVAQPOATQPTTLASVDSASHSVASGMDGEMV 840
 Db 721 QVLPQVASAGKOSTQVSVQVAPAEFVAVAQPOATQPTTLASVDSASHSVASGMDGEMV 780
 QY 841 PSSSGRHBGRTTKKHYRKSVSRSRHEKTSRPKILILVSNKGRVVECOLETNRKXVT 900
 Db 781 PSSSGRHBGRTTKKHYRKSVSRSRHEKTSRPKILILVSNKGRVVECOLETNRKXVT 840
 QY 901 FKPLDGNPREIKTIMNNDPILAIERESFVDQVREIIEKADEMLSDVSVEPEGDGL 960
 Db 841 FKPLDGNPREIKTIMNNDPILAIERESFVDQVREIIEKADEMLSDVSVEPEGDGL 900
 QY 961 ESLQKXDYGFSGSQKLEGEFKQPIPASMPQOIGIPTSSLTQVVSAGRRFIVSPVES 1020
 Db 901 ESLQKXDYGFSGSQKLEGEFKQPIPASMPQOIGIPTSSLTQVVSAGRRFIVSPVES 960
 QY 1021 RLRSKYFPESEITDTVAASTAQSPQMLSHSASSLSLQOAFSELRRAOMTEGENTAPNF 1080
 Db 961 RLRSKYFPESEITDTVAASTAQSPQMLSHSASSLSLQOAFSELRRAOMTEGENTAPNF 1020
 QY 1081 SHGTGTFPVVPPPLSSINGVPTTAAATAVPATSSPPNDISTSVIOSEVTVPTTEGLAGV 1140
 Db 1021 SHGTGTFPVVPPPLSSINGVPTTAAATAVPATSSPPNDISTSVIOSEVTVPTTEGLAGV 1080
 QY 1141 ATSTGVTVSGGLPIPVSESPVLSVSSITIPAVVSISSPQLQVPTSEIIVSSTA 1200
 Db 1081 ATSTGVTVSGGLPIPVSESPVLSVSSITIPAVVSISSPQLQVPTSEIIVSSTA 1140
 QY 1201 LYPSTVYATASAGGSTATPGPKRPVAVSQOAGSTTVGATLISVSTTTFPSTASQLS 1260
 Db 1141 LYPSTVYATASAGGSTATPGPKRPVAVSQOAGSTTVGATLISVSTTTFPSTASQLS 1200
 QY 1261 IOLSSSTSTPTLAETVVVVSASLSDKTSSTGGLAFSLSAPSSSSSSPAGVSSYSYOPRG 1320
 Db 1201 IOLSSSTSTPTLAETVVVVSASLSDKTSSTGGLAFSLSAPSSSSSSPAGVSSYSYOPRG 1260
 QY 1321 LHPVIPSIVASTPILPOAAGPTSTPILLPOVFSIPPLVQPVANVPVQOQLIHQOPAL 1380
 Db 1261 LHPVIPSIVASTPILPOAAGPTSTPILLPOVFSIPPLVQPVANVPVQOQLIHQOPAL 1320

Db 1233 LHPVIPSIVASTPILPOAAGPTSTPILLPOVFSIPPLVQPVANVPVQOQLIHQOPAL 1292
 QY 1381 LHPNPHHCPEVSDPOPKAPGIDDIKTLEEKI RSLSESHSSGAGHASYLERNSTYES 1440
 Db 1293 LHPNPHHCPEVSDPOPKAPGIDDIKTLEEKI RSLSESHSSGAGHASYLERNSTYES 1380
 QY 1441 TVTPTGIPPTVAPSKLSTSTSTCLPPTNLPLGVALPVVPVTPGVSPFVSSTTSQVK 1500
 Db 1381 TVTPTGIPPTVAPSKLSTSTSTCLPPTNLPLGVALPVVPVTPGVSPFVSSTTSQVK 1440
 QY 1501 PGTAPEKPLTKAPVLPVGTLPAGTLPSEQLPPEPSPSLTQSQOPIEDDUALRTLSF 1560
 Db 1441 PGTAPEKPLTKAPVLPVGTLPAGTLPSEQLPPEPSPSLTQSQOPIEDDUALRTLSF 1500
 QY 1561 EXTTTSAVGPVMAAATATTEAGTQOKVSOVQKEBPVLTATSGAVFMKRGROVVA 1620
 Db 1501 EXTTTSAVGPVMAAATATTEAGTQOKVSOVQKEBPVLTATSGAVFMKRGROVVA 1560
 QY 1621 DGAQKEGKNSKSEDAKSVHFESESTSESSVLSSESPSTLVKPEPAGITIPGISDVPEAH 1680
 Db 1561 DGAQKEGKNSKSEDAKSVHFESESTSESSVLSSESPSTLVKPEPAGITIPGISDVPEAH 1620
 QY 1681 KTTASEAKSDTQOPTKVGROVTTTANKVGRFVSXTEDKIDTQKKEGPAVASPPNDLEQ 1740
 Db 1621 KTTASEAKSDTQOPTKVGROVTTTANKVGRFVSXTEDKIDTQKKEGPAVASPPNDLEQ 1680
 QY 1741 AVLPVAVPKKEKPELSPSHLNGSPDEAPFLSRVDVDSGSPSHHOLSKSLPSONL 1800
 Db 1681 AVLPVAVPKKEKPELSPSHLNGSPDEAPFLSRVDVDSGSPSHHOLSKSLPSONL 1740
 QY 1801 SOSTLSNFSNYSNDNESEDIEDDLKLELRRLRDLKLEIODELOSKRHEIESLYTKLG 1860
 Db 1741 SOSTLSNFSNYSNDNESEDIEDDLKLELRRLRDLKLEIODELOSKRHEIESLYTKLG 1800
 QY 1861 KVPVAVIIPPAAPLSGRRRPTKSKGSSRSSSLGNKSPQLSGNLSGQSAASYLHPQOT 1920
 Db 1801 KVPVAVIIPPAAPLSGRRRPTKSKGSSRSSSLGNKSPQLSGNLSGQSAASYLHPQOT 1860
 QY 1921 LHPGNTPEGQONLQPLKPSPSDNLVAFSTSDGAI SVPSLSAPQGTSTNTVGTATV 1980
 Db 1861 LHPGNTPEGQONLQPLKPSPSDNLVAFSTSDGAI SVPSLSAPQGTSTNTVGTATV 1920
 QY 1981 NSQAQAQAPAMTSSRRGTFTDLHKLVDNWARDANMLSGRSGSGHMYTEGPMARXFS 2040
 Db 1921 NSQAQAQAPAMTSSRRGTFTDLHKLVDNWARDANMLSGRSGSGHMYTEGPMARXFS 1980
 QY 2041 APQQLCISMTSNLGSAPISAASATSLGHFTKSMCPQOQYGFPAITPGAGQSGTGCPAPQ 2100
 Db 2041 APQQLCISMTSNLGSAPISAASATSLGHFTKSMCPQOQYGFPAITPGAGQSGTGCPAPQ 2040
 QY 2101 PLGQFQVGTASLQNFNISNLQKSIINPQGSNLKTT 2136
 Db 2101 PLGQFQVGTASLQNFNISNLQKSIINPQGSNLKTT 2100
 QY 2136 PLGQFQVGTASLQNFNISNLQKSIINPQGSNLKTT 2048
 Db 2136 PLGQFQVGTASLQNFNISNLQKSIINPQGSNLKTT 2136

RESULT 14

ABO44408 standard; protein: 2141 AA.

ABO44408;

26-SEP-2003 (first entry)

Novel human protein kinase #28.

Human; kinase; enzyme; cosmetic application; nutraceutical application.

Homo sapiens.

US6541252-B1.

01-APR-2003.

XX 14-MAY-2001; 2001US-00854856.
 XX 19-MAY-2000; 2000US-0206015P.
 XX (LEXI-) LEXICON GENETICS INC.
 XX Walke DW, Hilburn E, Donoho G, Turner CA;
 XX MPI; 2003-575927/54.
 XX N-PSDB; ACH03791.
 XX
 XX New nucleic acid encoding novel human proteins, useful in cosmetic and
 XX nutritional applications.
 XX
 XX Disclosure; Page: 11pp; English.
 XX
 XX The invention relates to a new isolated nucleic acid encoding a novel
 XX human protein kinase. The nucleic acid is useful in cosmetic and
 XX nutritional applications. The present sequence represents the amino
 XX acid sequence of a novel human protein kinase. Note: The sequence data
 XX for this patent did not form part of the printed specification but was
 XX obtained in electronic format directly from USPTO at
 XX seqdata.uspto.gov/sequence.html?docid=654125281
 XX
 XX Sequence 2141 AA;
 XX
 XX Query Match 95.1%; Score 10287.5; DB 7; Length 2141;
 XX Best Local Similarity 94.4%; Pred. No. 0;
 XX Matches 2048; Conservative 0; Mismatches 0; Indels 121; Gaps 2;
 XX
 QY 61 MDKSRGAAATTTTTHRRFRFRSVICDSNATALELPGLEPLSIPQPSIPAAVQSAPEPEH 120
 Db 1 MDKSRGAAATTTTTHRRFRFRSVICDSNATALELPGLEPLSIPQPSIPAAVQSAPEPEH 60
 QY 121 REETVATATISQVAAQPPAAAPGBOAVAGPAPSTVPSTSKDRPVQSPLVSGSKEEPP 180
 Db 61 REETVATATISQVAAQPPAAAPGBOAVAGPAPSTVPSTSKDRPVQSPLVSGSKEEPP 120
 QY 181 ARSSGGGSAKEPPEBSQQDDIELEETKAVGMSNDRLFKPDIETGRSGFKTVYGLD 240
 Db 121 ARSSGGGSAKEPPEBSQQDDIELEETKAVGMSNDRLFKPDIETGRSGFKTVYGLD 180
 QY 241 TETTVAVAMCELQDRKLTSEKORFKBEAEMKGLQHPNIVRFYDSMESTYKGGKCIYV 300
 Db 181 TETTVAVAMCELQDRKLTSEKORFKBEAEMKGLQHPNIVRFYDSMESTYKGGKCIYV 240
 QY 301 TELMTSGTLKTYLKRFPYMKIKVLRSWCROILKGLQFLHTPTPIIHRDLKCDNIFITGP 360
 Db 241 TELMTSGTLKTYLKRFPYMKIKVLRSWCROILKGLQFLHTPTPIIHRDLKCDNIFITGP 300
 QY 361 TGSVKIGDLGLATLKRAFPKSVIGTPEFMAPEMYEKEKYDSVVVAFGCMLEMATSEY 420
 Db 301 TGSVKIGDLGLATLKRAFPKSVIGTPEFMAPEMYEKEKYDSVVVAFGCMLEMATSEY 360
 QY 421 PYSECQNAAOITRRVTSGVKRPASFDKVAIPEVKEIIEGCIRONKDERYSIKDLINHAFFQ 480
 Db 361 PYSECQNAAOITRRVTSGVKRPASFDKVAIPEVKEIIEGCIRONKDERYSIKDLINHAFFQ 420
 QY 481 EETGVARVELAEEDGEEKIATKLMRIDIKLKKGYKDNEMAIESFDLERDVPEDVAQEM 540
 Db 421 EETGVARVELAEEDGEEKIATKLMRIDIKLKKGYKDNEMAIESFDLERDVPEDVAQEM 480
 QY 541 VESGVYCEGDHKTAKAIKDRVSLIKRRBEOROLVREEOEKKKEESSIKQOVVQSSASQ 600
 Db 481 VESGVYCEGDHKTAKAIKDRVSLIKRRBEOROLVREEOEKKKEESSIKQOVVQSSASQ 540
 QY 601 TGIKOLPSASTGTIPJASTTSASVSTOVEPREPADQHOLOQYQOPSTISVLDGTVDSDGQ 660
 Db 541 TGIKOLPSASTGTIPJASTTSASVSTOVEPREPADQHOLOQYQOPSTISVLDGTVDSDGQ 600
 QY 661 SSVFTESRVSQOTVTSVGSQHEQASHSTGTVPGHIPSTVQAOSQPHGVYPPSSVAQGSQSG 720
 Db 661 SSVFTESRVSQOTVTSVGSQHEQASHSTGTVPGHIPSTVQAOSQPHGVYPPSSVAQGSQSG 720

Db 601 SSVFTESRVSQOTVTSVGSQHEQASHSTGTVPGHIPSTVQAOSQPHGVYPPSSV----- 653
 QY 721 QPSSSSLTGVSSSQPIQHPQQQGGIQQTAPPOQTVOYISLQSTSTSSSEATTAQVPSQPA 780
 Db 654 -----QQGICQQTAPPOQTVOYISLQSTSTSSSEATTAQVPSQPA 692
 QY 781 QVLPQVSAKQ----- 791
 Db 693 QVLPQVSAKQGFPEPLPQYFGDSNTAPSSNVASVCIHSTVYLKPEMTEVIATPGYFPT 752
 QY 792 -----STQGVSAVAPAPAV 807
 Db 753 VVQPYVESNLLVPMGVGVGVQVQVSGSLAQAPPTSSQAVLESTQGVSAVAPAPAV 812
 QY 808 AQPOATQPTTLASSVDSAHSDVVASGMDGNENPVSSGRHGGRTTKRHYKSVRSRSH 867
 Db 813 AQPOATQPTTLASSVDSAHSDVVASGMDGNENPVSSGRHGGRTTKRHYKSVRSRSH 872
 QY 868 KTSRPKRLINAVSNKGRVAVBCOLETHNRKMWTFKPDLDGNDPEIATIMVNNPILAI 927
 Db 873 KTSRPKRLINAVSNKGRVAVBCOLETHNRKMWTFKPDLDGNDPEIATIMVNNPILAI 932
 QY 928 RESFVDQVREIIIEKADWLSQVSEVPEEGDGLSLQKDDYGRSGSKLEGEKOPIPA 987
 Db 933 RESFVDQVREIIIEKADWLSQVSEVPEEGDGLSLQKDDYGRSGSKLEGEKOPIPA 992
 QY 988 SSMPOQIGIPSSLTQVHSGRRFIVSPVRESLRESKVPSEITTDVAASTQSPGM 1047
 Db 993 SSMPOQIGIPSSLTQVHSGRRFIVSPVRESLRESKVPSEITTDVAASTQSPGM 1052
 QY 1048 LSHSASSLSLQQAASELRRAQMTGPTAPNFGHTGTFPVVPPFLSSIAIGVPTTAAT 1107
 Db 1053 LSHSASSLSLQQAASELRRAQMTGPTAPNFGHTGTFPVVPPFLSSIAIGVPTTAAT 1112
 QY 1108 APVATSSPNDISTSVIOSEVIVPTEEGIAGVATSGVTSGGLPVPVSESVLSVV 1167
 Db 1113 APVATSSPNDISTSVIOSEVIVPTEEGIAGVATSGVTSGGLPVPVSESVLSVV 1172
 QY 1168 SSTTPAVVSISSTPSLQVPTSTSEIVSTALIPSVTASATASAGSIAATGPPRPA 1227
 Db 1173 SSTTPAVVSISSTPSLQVPTSTSEIVSTALIPSVTASATASAGSIAATGPPRPA 1232
 QY 1228 VVSQQAAGSTTVGATLTSVSTTSFPTSTASQLSTLSSSTPTLAEVTVVVASLDTTS 1287
 Db 1233 VVSQQAAGSTTVGATLTSVSTTSFPTSTASQLSTLSSSTPTLAEVTVVVASLDTTS 1292
 QY 1288 HSSTTGLAFSLASBSSSSPGAGVSYISQPGHLPVIBSVIASTPILPQAAGPTSTPL 1347
 Db 1293 HSSTTGLAFSLASBSSSSPGAGVSYISQPGHLPVIBSVIASTPILPQAAGPTSTPL 1352
 QY 1348 LPQVPSIPPLVOPVANYPAVOQTLIHSQPOPALPNPHTHCPEVVDSTQPKAGIDIK 1407
 Db 1353 LPQVPSIPPLVOPVANYPAVOQTLIHSQPOPALPNPHTHCPEVVDSTQPKAGIDIK 1412
 QY 1408 TLEBKLSLSEHSSSGAQSASVLETSVLESTVGTGIPTTAAVAPSKLITSTSTCLP 1467
 Db 1413 TLEBKLSLSEHSSSGAQSASVLETSVLESTVGTGIPTTAAVAPSKLITSTSTCLP 1472
 QY 1468 TNLPLGTVALPVTVPVTPGVSTPSTTSGVKGPINAPSKPPLTKAPVLPVTELPAGTL 1527
 Db 1473 TNLPLGTVALPVTVPVTPGVSTPSTTSGVKGPINAPSKPPLTKAPVLPVTELPAGTL 1532
 QY 1528 PSEQLPPEPGPSLTQSQOPLLEDLDAQRLTSLPEXITVTSAVGVGVAAPALTAATEATOP 1587
 Db 1533 PSEQLPPEPGPSLTQSQOPLLEDLDAQRLTSLPEXITVTSAVGVGVAAPALTAATEATOP 1592
 QY 1588 QKGVSYKESPVLATSSGAVFKGRFQVVAADGAKESKNSSEDAKSHFESSSTESS 1647
 Db 1593 QKGVSYKESPVLATSSGAVFKGRFQVVAADGAKESKNSSEDAKSHFESSSTESS 1652
 QY 1648 VLSSSSESTLVKPEPNTGITPGIISDVPESAHKTASAKSDTQPTKYVGRPOVTTAN 1707
 Db 1653 VLSSSSESTLVKPEPNTGITPGIISDVPESAHKTASAKSDTQPTKYVGRPOVTTAN 1712

QY 792 -----STQGVQVAPAEPAVVAQOATOPTTLASVDYSAHSDVAGS 834
 Db 993 GSLAQPTTSSQOAVLESTQGVQVAPAEPAVVAQOATOPTTLASVDYSAHSDVAGS 992
 QY 895 DGENVBSGGHHERTTKRHRKSVRSRREKTSRPRLRLVNSKGRVVECOLETH 894
 Db 993 DGENVBSGGHHERTTKRHRKSVRSRREKTSRPRLRLVNSKGRVVECOLETH 1052
 QY 895 NRKMTFFEDLDGNDPEEIIATIMVANDFLIATERESFVQVVEIEKXDEMISEVSVSEP 954
 Db 1053 NRKMTFFEDLDGNDPEEIIATIMVANDFLIATERESFVQVVEIEKXDEMISEVSVSEP 1112
 QY 955 EGDQGLSLQCKDQVGFSGSQKLEGEFKQPIPASSMPOQIGIPSSLTQVHSAGRFIV 1014
 Db 1113 EGDQGLSLQCKDQVGFSGSQKLEGEFKQPIPASSMPOQIGIPSSLTQVHSAGRFIV 1172
 QY 1015 SPVPSRLRESKVFSEITTDVYAATAOSPGMNLSSHSSLSLQOAFSELRAQMTSEPN 1074
 Db 1173 SPVPSRLRESKVFSEITTDVYAATAOSPGMNLSSHSSLSLQOAFSELRAQMTSEPN 1232
 QY 1075 TAPNFSHTGTFPVVPPFLSSIAGVPTTAATAVAPATSSPNDISTSVIOSEVTPTE 1134
 Db 1233 TAPNFSHTGTFPVVPPFLSSIAGVPTTAATAVAPATSSPNDISTSVIOSEVTPTE 1292
 QY 1135 EGIAGVATSTGVVTSQGLPIPVSESPVLSVSSITIPAVVSIITTSPLQVPTSTSEI 1194
 Db 1293 EGIAGVATSTGVVTSQGLPIPVSESPVLSVSSITIPAVVSIITTSPLQVPTSTSEI 1352
 QY 1195 VVSSTALYPSVTGATSSAGSGSTATBPKPPAVVSOQAAGSTTVGATLTSTVSTTSFPS 1254
 Db 1353 VVSSTALYPSVTGATSSAGSGSTATBPKPPAVVSOQAAGSTTVGATLTSTVSTTSFPS 1412
 QY 1255 TASQSLQSLSSSTSTPTLAETVWVSAHSLDKTSHSTGLAFSLJAPSSSSSPGAGVSSY 1314
 Db 1413 TASQSLQSLSSSTSTPTLAETVWVSAHSLDKTSHSTGLAFSLJAPSSSSSPGAGVSSY 1472
 QY 1315 ISQPGHLPVTPSVIASTPILPQAAGTSTPLLPVPSIPPLVQVAVNPVAVOQTLIHS 1374
 Db 1473 ISQPGHLPVTPSVIASTPILPQAAGTSTPLLPVPSIPPLVQVAVNPVAVOQTLIHS 1532
 QY 1375 QOPALLPNQHPTCEPVDSTQPKAPGIDIKILEEKLRSLSFSEHSSSGAQNASVLEI 1434
 Db 1533 QOPALLPNQHPTCEPVDSTQPKAPGIDIKILEEKLRSLSFSEHSSSGAQNASVLEI 1592
 QY 1435 SLVSESTVTPGIPTTAVAPSKLTSTSTCLPPTNLPGLTALPVTVPVTPGOVSTPST 1494
 Db 1593 SLVSESTVTPGIPTTAVAPSKLTSTSTCLPPTNLPGLTALPVTVPVTPGOVSTPST 1652
 QY 1495 TTSGVKFGTAPSKPPLTKAPVLPVGTLPAGTLPSEQLPFPGPSLTQSQOPLLEDLAOL 1554
 Db 1653 TTSGVKFGTAPSKPPLTKAPVLPVGTLPAGTLPSEQLPFPGPSLTQSQOPLLEDLAOL 1712
 QY 1555 RRTLSPEXITVTSVAVGVMAAPATITBAGTQPKGVQVKEGVLATSSGAGVFKMGRF 1614
 Db 1713 RRTLSPEXITVTSVAVGVMAAPATITBAGTQPKGVQVKEGVLATSSGAGVFKMGRF 1772
 QY 1615 QVSVAADGAQKEGKNKSEDAKSVFSESTSESSVLSSSSPSTLWKPEPNGITIPGISD 1674
 Db 1773 QVSVAADGAQKEGKNKSEDAKSVFSESTSESSVLSSSSPSTLWKPEPNGITIPGISD 1832
 QY 1675 VPESAHKTTASEAKSDTGQPTKVGRFOVTTTANKVGRFVSXKTEKITDITKKEGPVAPSP 1734
 Db 1833 VPESAHKTTASEAKSDTGQPTKVGRFOVTTTANKVGRFVSXKTEKITDITKKEGPVAPSP 1892
 QY 1735 FMDLEQAVLPAVLPKKEKPELSESHNGRSDPEAFLSRDUDGSGSPHSHQLSKS 1794
 Db 1893 FMDLEQAVLPAVLPKKEKPELSESHNGRSDPEAFLSRDUDGSGSPHSHQLSKS 1952
 QY 1795 LPSQNTLSQSLNSFNSSYMSDNEJEDBDLKLRLRDKHLKEIODLOSROKHEIES 1854
 Db 1953 LPSQNTLSQSLNSFNSSYMSDNEJEDBDLKLRLRDKHLKEIODLOSROKHEIES 2012
 QY 1855 LYTTLGKVPVAVIIPPAAPLSGRRRPTKSGSKSRSSSLGNKSPQLSGNTLSQGSAAV 1914

Db 2013 LYTTLGKVPVAVIIPPAAPLSGRRRPTKSGSKSRSSSLGNKSPQLSGNTLSQGSAAV 2072
 QY 1915 LHPQOTLHPNIPESGONOLQPLKPSSSDNLYSAFTSDGALSVPLSAPGQGTSSSTN 1974
 Db 2073 LHPQOTLHPNIPESGONOLQPLKPSSSDNLYSAFTSDGALSVPLSAPGQGTSSSTN 2132
 QY 1975 TVGATVNSQAAQAPPAATSSRKGTFTDHLKLVDMNARDAMNLSGRRGSKGHNNYEGPG 2034
 Db 2133 TVGATVNSQAAQAPPAATSSRKGTFTDHLKLVDMNARDAMNLSGRRGSKGHNNYEGPG 2192
 QY 2035 MARKFSAPGOLCISNTSNLGGSAPIASASATSLGHFTSMCPPOQYGPAPFPGAOWSGT 2094
 Db 2193 MARKFSAPGOLCISNTSNLGGSAPIASASATSLGHFTSMCPPOQYGPAPFPGAOWSGT 2252
 QY 2095 GGPAPQPLGQFQPVGTASLQNFNINSLQKSIINPSPGSLRTT 2136
 Db 2253 GGPAPQPLGQFQPVGTASLQNFNINSLQKSIINPSPGSLRTT 2294

Search completed: September 24, 2004, 01:06:44
 Job time : 213 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: September 25, 2004, 05:24:04 ; Search time 1217 Seconds
(without alignments)
7380.239 Million cell updates/sec

Title: US-10-010-720-14
Perfect score: 10812
Sequence: 1 MSGAARXKQSTPGSLFLSP.....NISNLRKSIISNPPGSNLRRT 2136

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 6597448 seqs, 2102469792 residues
Total number of hits satisfying chosen parameters: 13194896

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-O=/cg2_1/USPRO_spool/US10010720/runat_23092004_154321_25331/app_query.fasta.1.2311
-DB=Pending_Patents_NA_New -QFMT=fastap -SUFFIX=rpna -MINMATCH=0.1 -LOOCPD=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdt
-LIST=45 -DOCALLIGN=200 -THR SCORE=PCT -THR MAX=100 -THR MIN=0 -ALIGN=15
-MOB=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10010720 @CGN 1.1 249 @runat 23092004 154321 25331 -NCPU=6 -ICPU=3
-NO MMAP -LARGEIOBRY -NEG_SCORES=0 -WAIT -DSPELOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -MAIN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

1: Pending Patents NA New:*
2: /cg2_6/ptodata/1/pna/PCT_NEW_COMB.seq:*
3: /cg2_6/ptodata/1/pna/US07_NEW_COMB.seq:*
4: /cg2_6/ptodata/1/pna/US08_NEW_COMB.seq:*
5: /cg2_6/ptodata/1/pna/US09_NEW_COMB.seq:*
6: /cg2_6/ptodata/1/pna/US10_NEW_COMB.seq:*
7: /cg2_6/ptodata/1/pna/US10_NEW_COMB.seq2:*
8: /cg2_6/ptodata/1/pna/US10_NEW_COMB.seq3:*
9: /cg2_6/ptodata/1/pna/US60_NEW_COMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	10677	98.8	1	PCT-US04-24424-2014
2	10677	98.8	1	PCT-US04-27459-1
3	10646	98.5	7	US-10-491-467-54
4	10604	98.1	8678	7 PCT-US04-02652-450
5	10604	98.1	8678	7 US-10-767-471-450
6	10604	98.1	10464	1 PCT-US04-02652-455
7	10604	98.1	10464	7 US-10-767-471-455
8	10567.5	97.7	7853	1 PCT-US04-02652-452
9	10567.5	97.7	7853	7 US-10-767-471-452

10	10567.5	97.7	9639	1	PCT-US04-02652-454	Sequence 454, App
11	10567.5	97.7	9639	7	US-10-767-471-454	Sequence 454, App
12	9737	90.1	9417	1	PCT-US04-02652-449	Sequence 449, App
13	9737	90.1	9417	7	US-10-767-471-449	Sequence 449, App
14	9729.5	90.0	7328	5	US-09-979-167-63	Sequence 63, App
15	7154	66.2	4872	1	PCT-US04-02652-453	Sequence 453, App
16	7154	66.2	4872	7	US-10-767-471-453	Sequence 453, App
17	5931	54.9	5785	7	US-10-021-6984-3606	Sequence 3606, App
18	5902	54.6	6294	1	PCT-US04-02652-451	Sequence 451, App
19	5902	54.6	6294	7	US-10-767-471-451	Sequence 451, App
20	3132.5	29.0	170546	1	PCT-US04-02652-10703	Sequence 10703, A
21	3132.5	29.0	170546	7	US-10-767-471-10703	Sequence 10703, A
22	3026	28.0	1835	6	US-10-765-700-120	Sequence 120, App
23	2719.5	25.2	5637	6	US-10-932-349-491	Sequence 491, App
24	2642	24.4	7825	7	US-10-490-592-6	Sequence 4, App
25	2639.5	24.4	6582	7	US-10-490-592-6	Sequence 6, App
26	2638.5	24.4	6891	1	PCT-US04-27459-5	Sequence 5, App
27	2627	24.3	1536	5	US-09-552-317-7020	Sequence 7020, App
28	2623	24.3	6834	1	PCT-US04-27459-3	Sequence 3, App
29	2526.5	23.4	7569	6	US-10-932-349-494	Sequence 494, App
30	2466	22.8	5403	1	PCT-US04-27459-6	Sequence 6, App
31	2441.5	22.6	7278	1	PCT-US04-27459-8	Sequence 8, App
32	2355.5	21.6	6807	1	PCT-US04-27459-7	Sequence 7, App
33	2130	19.7	3861	1	PCT-US04-27459-9	Sequence 9, App
34	2096.5	19.4	4064	6	US-10-932-349-356	Sequence 356, App
35	2096.5	19.4	4341	6	US-10-932-349-357	Sequence 357, App
36	1974	18.3	1928	1	PCT-US04-27459-2	Sequence 2, App
37	1922	17.8	1345	7	US-10-021-6984-3609	Sequence 3609, App
38	1807	16.7	2404	1	PCT-US04-27459-4	Sequence 4, App
39	1802	16.7	2328	6	US-10-932-349-493	Sequence 493, App
40	1718	15.9	1667	7	US-10-100-683-2951	Sequence 2951, App
41	1508	13.9	1179	7	US-10-021-6984-3607	Sequence 3607, App
42	1444	13.4	893	6	US-10-472-965-141	Sequence 141, App
43	1444	13.4	893	6	US-10-896-164-1520	Sequence 1520, App
44	1444	13.4	893	6	US-10-868-194B-1520	Sequence 1520, App
45	1444	13.4	893	6	US-10-868-184B-1520	Sequence 1520, App

ALIGNMENTS

RESULT 1
PCT-US04-24424-2014
; Sequence 2014, Application PC/TUS0424424
; GENERAL INFORMATION:
; APPLICANT: Li, Martha
; APPLICANT: Rudnow, Brent A.
; APPLICANT: Webster, Kevin R.
; APPLICANT: Jackson, Donald
; APPLICANT: Wong, Tai W.
; TITLE OF INVENTION: BIOMARKERS OF CYCLIN-DEPENDENT KINASE MODULATION
; FILE REFERENCE: D0310 PCT
; CURRENT APPLICATION NUMBER: PCT/US04/24424
; PRIOR APPLICATION NUMBER: 2004-08-13
; PRIOR FILING DATE: 2003-07-29
; NUMBER OF SEQ ID NOS: 2786
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2014
; LENGTH: 7149
; TYPE: DNA
; ORGANISM: Homo sapiens
PCT-US04-24424-2014

Alignment Scores:

Pred. No.: 0
Score: 10677.00
Percent Similarity: 89.63%
Best Local Similarity: 89.63%
Query Match: 98.75%
DB: 1
Gaps: 1

US-10-010-720-14 (1-2136) x PCT-US04-24424-2014 (1-7149)
Length: 7149
Matches: 2135
Conservative: 0
Mismatch: 1
Indels: 246

```

QY 1 MetSerGlyGlyAlaIaGluLeuSerSerThrProGlySerLeuPheLeuSerPro 20
Db 1 ATGCTGGCGCGCGCGAGAGAGACAGACACTCCCGGTTCCCTGTTCTCTCGCGG 60
QY 21 ProAlaProAlaProLysAsnGlySerSerSerSerSerValGlyGlyLeuGly 40
Db 61 CCGGCTCTGCGCGCGAGAGATGGCTCCGATCTCCGATCTCCGCGGGAGAACTGGGA 120
QY 41 AlaAlaAlaAlaAspAlaValThrGlyArgThrGluGluTyrArgArgArgHisThr 60
Db 121 GCCCGCGCGCGCGAGCTGTGACCGGACGAGACGAGAGTACAGCGCGCGCGCACACT 180
QY 61 MetAspLysAspSerSerArgLysAlaAlaAlaThrThrThrThrThrGluHisArgPhePhe 80
Db 181 ATGACACAGACAGACCGTGGGGCGCGCGCGCACACTACACACTGACACCGCTTCTTC 240
QY 81 ArgArgSerValLysAspSerSerAsnAlaThrAlaLeuGluLeuProGlyLeuProLeu 100
Db 241 CGCGGAGCGCTCATCTGCGACTCCCAATGCCACTGCGTGGAGCTTCCGCGGCTTCTCTT 300
QY 101 SerLeuProGlnProSerLysProAlaAlaValProGlnSerAlaProProGlnProHis 120
Db 301 TCCCTGCGCGCGCGAGATCCCGCGGCTGTCCGACAGTGTCTCCACCGAGCGCGCAC 360
QY 121 ArgGluGluThrValThrAlaThrAlaThrSerGlnValAlaGlnGlnProProAlaAla 140
Db 361 CGGAGAGAGACCGTGAACCGCCACCTTCCAGGTAGCCGACGACCGCTCCAGCGCT 420
QY 141 AlaAlaProGlyGluGlnAlaValAlaGlyProAlaProSerThrValProSerSerThr 160
Db 421 GCCCGCGCGCGAGACGCGCTGCGCGGCGCTCCCTCGACTGTCGCCAGAGACC 480
QY 161 SerLysAspArgProValSerGlnProSerLeuValGlySerLysGluGluProProPro 180
Db 481 AGCAAGACGCGCGAGTCCAGCTTGTGGGAGCAAGAGAGAGAGAGAGAGAGAGAGAG 540
QY 181 AlaArgSerGlySerGlyGlySerAlaLysGluProGlnGluGluArgSerGlnGln 200
Db 541 GCGAGAGTGGCAACGCGCGCGCGAGCGCCACAGACGACGAGAGAGAGAGAGAGAGAG 600
QY 201 GlnAspAspLysGluGluLeuGluThrLysAlaValGlyMetSerAsnAspGlyArgPhe 220
Db 601 CAGATATATATCGAAGAGCTGAGACCAAGCGCTGGGAGATGTTAAAGATGCGCGCTT 660
QY 221 LeuLysPheAspLysGluLysArgGlySerPheLysThrValTyrLysGlyLeuAsp 240
Db 661 CTCAAGTTTGACATGGAATCGGAGAGGCTCTTTAAGACGCTCAAAAGGCTCTGAC 720
QY 241 ThrGluThrThrValGluValAlaTyrCysGluLeuGlnAspArgLysLeuThrLysSer 260
Db 721 ACTGAAACCAACCGTGAAGTGGCTGGTGTGAATGACAGATCGAAAATTAAACAAAGTCT 780
QY 261 GlnArgGlnArgPheLysGluGluAlaGluMetLeuLysGlyLeuGlnHisProAsnLys 280
Db 781 GAGAGGAGAGATTTAAAGAAAGAGCTGAATGTTAAAGAGTCTTCGACATCCCAATATT 840
QY 281 ValArgPheTyrAspSerLysThrGluSerThrValLysGlyLysLysCysLysLeuVal 300
Db 841 GTTAGATTTTATGATTTCTGGGAATCCAGATAAAGAAAGAAAGATGATGTTTGGTG 900
QY 301 ThrGluLeuMetThrSerGlyThrLeuLysThrTyrLeuLysArgPheLysValMetLys 320
Db 901 ACTGAACTTTATGACGTCTGGAACACTTAAACGATCTGAAAAGCTTTAAAGATGAAG 960
QY 321 LysLysValLeuArgSerTyrCysArgGlnLysLeuLysGlyLeuGlnPheLeuHisThr 340
Db 961 ATCAAAATTTCAAAAGCTGTGCGCTGAGATCTTTAAAGGTCTTCGATTTCTTCATACT 1020
QY 341 ArgThrProProLysLysHisArgAspLeuLysCysAspAsnLysPheLysThrGlyPro 360
Db 1021 CGAATCTACCTATCATTCACCGGAGCTTAAAGTAAACAACTTTATTCACCGGCGCT 1080
QY 361 ThrGlySerValLysLysGlyAspLeuGlyLeuAlaThrLeuLysArgAlaSerPheAla 380

```

```

Db 1081 ACTGCTCAGTCAGATGAGAGACTCGTCTGGCAACCTTGAAAGCGGCTTCTTTGCC 1140
QY 381 LysSerValLysGlyThrProGluPheMetAlaProGluMetTyrGluGluLysTyrAsp 400
Db 1141 AAGAGTGTAGTACGATCCCGAGAGTTCATGGCCCTTGAGATGTATGAGAGAAATATGAT 1200
QY 401 GluSerValAspValTyrAlaPheGlyMetCysMetLeuGluMetAlaThrSerGlyTyr 420
Db 1201 GAATCCGTGTACGTTTATGCTTTGGATGTGATGATGATGATGATGATGATGATGAT 1260
QY 421 ProTyrSerGluCysGlnAsnAlaAlaGlnLysTyrArgArgValThrSerGlyValLys 440
Db 1261 CTTACTCGAGAGTCCAAATATGCTGCGCAGATCTACCGTGGCGTACGACAGTGGGTGAG 1320
QY 441 ProAlaSerPheAspLysValAlaAlaLeuProGlnValLysGluLysGluLysGlyCysLys 460
Db 1321 CCAGCCAGTTTGTCAAAAGTATGCAATTCCTGAAGTGAAGAAATTTAAGAGATGACATA 1380
QY 461 ArgGlnAsnLysAspGluArgTyrSerLysLysAspLeuLeuAsnHisAlaPhePheGln 480
Db 1381 CGACAAACAAAGAGTAAAGATATTCATCAAAACCTTTTGAACCATGCTTCTTCCAA 1440
QY 481 GlnGluThrGlyValArgValGluLeuAlaGluGluAspAspGlyLysLysLysLysLys 500
Db 1441 GAGGAAACAGAGAGTACCGGTATGAAATTTAGCAGAGAAATGATGAGAGAAAAATAGCCATA 1500
QY 501 LysLeuThrLeuArgLysGluAspLysLysLysLysLysLysLysLysLysLysLysLys 520
Db 1501 AAATATAGCTACGATATTTGAAGATATTAAGAAATTTAAGAGAAATTTAAGAAATTTAAGAA 1560
QY 521 AlaLysLysPheSerPheAspLeuGluArgAspValProGluAspValAlaGlnGluMet 540
Db 1561 GCTATTAGTCTTCTTTTGTATTTAGAGAGAGATCCAGAGAGTGTGCAAGAAATG 1620
QY 541 ValGluSerGlyTyrValCysGluGlyAspHisLysThrMetAlaLysAlaLysAsp 560
Db 1621 GTTAGATCTGGTATGTCTGTGAAGGTGATCACAAGACATGGTTAAAGCTATTAAGAC 1680
QY 561 ArgValSerLeuLysArgLysArgGluGlnArgGluLeuValArgGluGluGln 580
Db 1681 AGAGTATCATTTAATTAAGAGAAACGAGACAGCGGAGTGTGTAACGAGAGAAACAA 1740
QY 581 LysLysLysGlnGluGluSerSerLeuLysGlnGlnAlaGluGlnSerSerAlaSerGln 600
Db 1741 AAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1800
QY 601 ThrGlyLysGluLeuProSerAlaSerThrGlyLysProThrAlaSerThrThrSer 620
Db 1801 ACGAGATCAAGACGCTCCCTTGTGTAACCGGACATCCTACTGCTTACCACTTCA 1860
QY 621 AlaSerValSerThrGlnValGluProGluGluProGluAlaAspGlnHisGlnGln 640
Db 1861 GCTTCACTTCTACACAAAGTAAACCTGAAGACCTAGGACAGATCAACATCAACACTA 1920
QY 641 GlnTyrGlnGlnProSerLysSerValLeuSerAspGlyThrValAspSerGlyGlnGly 660
Db 1921 CAGTACAGCAACCCCAATATATCTGTATCTGATGGAGCGGTTGACAGTGTGCAAGGA 1980
QY 661 SerSerValPheThrGluSerArgValSerSerGlnGlnThrValSerTyrGlySerGln 680
Db 1981 TCCGTGCTTTCACAGATCTCGAGTGAAGCAGCAACAGACAGTTTCAATATGTTCCCA 2040
QY 681 HisGluGlnAlaHisSerThrGlyThrValProGlyHisLysProSerThrValGlnAla 700
Db 2041 CATGAACAGCACATTTCTACAGGACAGTCCAGGCGATTAACCTTCTACCTGTCACAGCA 2100
QY 701 GlnSerGlnProHisGlyValTyrProProSerSerValAlaGlnGlyGlnSerGlnGly 720
Db 2101 CAGCTTACGCCCAATGGGCTATATTCACCTTCAAGTGTGGACAGGGGCGAGGCCAGG 2160
QY 721 GlnProSerSerSerSerLeuThrGlyValSerSerSerGlnProLysGlnHisProGln 740

```

Dp	2161	AAGCATCTCTCAAGTAGCTTAACAGGGGTTTCATCTTCCCAACCATACACATCTCCAG	2222
Qy	741	GIingIingIyIleGIingInThraIaProProGIingInThraIaGIInTySerIeuSer	760
Dp	2221	CAGCAGCAGGGGAATACAGACAGACACCCCTCCTCAACAGACAGTSCATATTCACCTTCA	2288
Qy	761	GIInThrSerThSerSerGIuIaIaThrThraIaGIInProValSerGIInProGIInIaPro	780
Dp	2281	CAGACATCAACTCCAGTAGAGGCCACTACTGCACGCCAGTAGAGTCAACCTTCAGCTCCA	2340
Qy	781	GIInValIeuProGIInValSerIaGIyIyGIIn-----	791
Dp	2341	CAAGCTTCCCTCAAGTATCACTGGAAAAACGCTTCCAGTTTCCACCCAGTACCACT	2400
Qy	791	-----	791
Dp	2401	ATCCAAGCGAAGCTCAGATCCAGTTGGACACAAACCTGGTTGTTCCAATCCACTCT	2466
Qy	791	-----	791
Dp	2461	GGTGGCTATTTCTTCCAGTGGGACAGCGCTCCCTACTCCCTTGGCTCCAGTACCCT	2520
Qy	791	-----	791
Dp	2521	GTCTCTCAGATTCCCATATCAACTCTCATGTGTTCTAGGGCTCAGACAGGTTTCTCATCC	2580
Qy	791	-----	791
Dp	2581	CTTCCCATCACAATGAGCAGCTGGCATTTACTCAGCCTGTGTCAGTTGGCTTATCTGCT	2640
Qy	791	-----	791
Dp	2641	ACAACAGCTGCGATCCCGGGGGTATCAACTGTGTTCCTAGTCAAGCTTCCAACCTTCTCG	2700
Qy	791	-----	791
Dp	2701	CAGCCTGTGACTCAGCTGCCAAATCAGTTCAACCAAGCTCTACAAACAGCAGTTCCAG	2766
Qy	791	-----	791
Dp	2761	TCCATGGGAATACACAGTCACTTGGACAGAGTGCTGAGTTCCACTTCTCTGGAGAT	2820
Qy	791	-----	791
Dp	2821	GTTCTGTACAGGGCTTCCACCTCAGCTGCACACACAGTACCAGAGATTCAAAATATT	2880
Qy	791	-----	791
Dp	2881	GCTCCCTCTTCCACGTGGCTTGTGTTGCATCTTAAGATCTTATCCCTCCCATG	2940
Qy	791	-----	791
Dp	2941	CCGACAGAGTAAGTACCTGCTTACACTGGGTAATTTCACAGTGTGTCACACCTTATGTGAA	3000
Qy	791	-----	791
Dp	3001	TCAATCTTTTAGTCTTATGGGTGGTGTAGAGAGCAGGTTCAAGTCCACGCAGGA	3066
Qy	792	-----SerThGIIn	794
Dp	3061	GGAGATTACACAAAGCCCCCATACATACCTCCACGACAGCATTTTGGAGATTAATCAG	3120
Qy	795	GIyValSerGIInValIaIaProIaGIuProValIaIaValIaGIInProGIInIaIaThrGIIn	814
Dp	3121	GGAGCTCTTCAGGTGCTCTGCGAGAGCAGTGTGATAGCAGAGCCCAAGCTACCCAG	3180
Qy	815	ProThrThrIeuIaIaSerSerValaIaIaSerIaIaIaSerIaIaIaSerGIyIeSer	834
Dp	3181	CCGACCACTTTGGCTTCTCTGTGAGACAGTGACATTCAGATGTTGCTTCAGATATGAGT	3240
Qy	835	AspGIyIaIaIaSerIaIaProSerSerSerGIyIaIaIaGIuIaIaIaThrThrIaIaIa	854
Dp	3241	GATGGCAATGAGAAAGTCCCATCTTCCAGTGGAAAGCATGAAAGGAAGACTACAAACGG	3300

[illegible]

Db	6541	CTGTTACAGGCCCTTAAAGCATCTCCCTCCAGTGAACAACCTCTATTGACGCTTCACAGT	6600
Qy	1955	AspGlyAlaIleSerValProSerLeuSerAlaProGlyGlnGlyThrSerSerThrAsn	1974
Db	6601	GATGGTGCATTCTAGTACCAAGCCTTCTCTCCAGGTCAAGAAACACGACGACAAAC	6660
Qy	1975	ThrValGlyAlaThrValAsnSerGlnAlaIleGlnIleGlnIleProProAlaMetThrSer	1994
Db	6661	ACTGTGGGGGACACGTAACAGCAAGCCACCCGCAAGTCAAGCCTCTGACCATATACGTC	6720
Qy	1995	SerArgLysGlyThrPheThrAspAspLeuHisLysLeuValAspAsnTrpAlaArgAsp	2014
Db	6721	AGCAGGAAGGCACTTTCACAGATGACTTGCACAAGTTGTTAGCAATTGGGCCGAGAT	6780
Qy	2015	AlaMetAsnLeuSerGlyValArgArgLysSerLysIleMetAsnTrpGlnGlyProGly	2034
Db	6781	GCCATGATCTCTCAGGACGAGAGGAAAGGAAGGCAACATGAATTACGAAGGGCCCTGGA	6840
Qy	2035	MetLalaArgLysPheSerAlaProGlyGlnLeuCysIleSerMetThrSerAsnLeuGly	2054
Db	6841	ATGGCAGGAAGTTCTCTGCACCTGGGCAACTGTGCATCTCCATGACCTCCAACTGGGT	6900
Qy	2055	GlySerAlaProIleSerAlaAlaSerAlaThrSerLeuGlyHisPheThrLysSerMet	2074
Db	6901	GGCTCTGCCCCCATCTCTGCAGCATCACTCACTCTCAAGGTCACTTCACCAAGACTATG	6960
Qy	2075	CysProProGlnGlnTrpGlyPheProAlaThrProPheGlyValIleGlnTrpSerGlyThr	2094
Db	6961	TGCCCCCACAGCATATAGGCTTTCACACTACCCCATTTGGGGCTCAATGAGTGGGACG	7020
Qy	2095	GlyGlyProAlaProGlnProLeuGlyGlnPheGlnProValGlyThrAlaSerLeuGln	2114
Db	7021	GGTGGCCCAAGCACCAAGCACTTGGCCAGTTCCAACTGTGGGAACCTGCTCTTGACG	7080
Qy	2115	AsnPheAsnIleSerAsnLeuGlnLysSerIleSerAsnProProGlySerAsnLeuArg	2134
Db	7081	AATTTCACATCAGCATTTTGCAGAAATCCATCACAACAACCCCCAGGCTCCCAACTGCGG	7140
Qy	2135	ThrThr2136	
Db	7141	ACCACT7146	

RESULT 2
PCT-US04-27459-1
; Sequence 1, Application PC/TUS0427459

```

/ GENERAL INFORMATION:
/ APPLICANT: EXELIXIS, INC.
/ TITLE OF INVENTION: PREKMTNS AS MODIFIERS OF THE RAC PATHWAY AND METHODS OF T
/ FILE REFERENCE: EX04-060C-PC
/ CURRENT APPLICATION NUMBER: PCT/US04/27459
/ CURRENT FILING DATE: 2004-08-30
/ PRIOR APPLICATION NUMBER: US 60/495,193
/ PRIOR FILING DATE: 2003-08-14
/ NUMBER OF SEQ ID NOS: 13
/ SOFTWARE: PatentIn version 3.2
/ SEQ ID NO 1
/ LENGTH: 7149
/ TYPE: DNA
/ ORGANISM: Homo sapiens
PCT-US04-27459-1

Alignment Scores:
Pred. No.: 0
Score: 10677.00
Percent Similarity: 89.63%
Best Local Similarity: 89.63%
Query Match: 98.75%
DB: 1
Gaps: 1
Length: 7149
Matches: 2135
Conservative: 0
Mismatches: 1
Indels: 246
Gaps: 1

```

QY 381 LysSerValIleGlyThrProGluPheMetAlaProGluMetTyrGluGluLysTyrAsp 400
 Db 1141 AAGAGTGTGATAGTACCCAGAGTTCATGGCCCTGAGATGTATGAGAGAGAAATATCAT 1200
 QY 401 GluSerValAspValTyrAlaPheGlyMetCysMetLeuGluMetAlaThrSerGluTyr 420
 Db 1201 GAATCCGTTGACGTTATGCTTGGAGTGTGATGCTTGAGATGGCTACATCTGAATAT 1260
 QY 421 ProTyrSerGluCysGluAsnAlaIleGlnIleTyrArgValThrSerGluValLys 440
 Db 1261 CTTTACTGGAGTGCCTCAAAATGCTGCGCATCTACCGTCCGGTGAACAGTGGGTGAG 1320
 QY 441 ProAlaSerPheAspLysValAlaIleProGluValLysGluIleIleGluGlyCysIle 460
 Db 1321 CCAGCCAGTTTGTGCAAGTATGCAATTCCTGAGTGAAGAAATATTATGAGAGATGCTA 1380
 QY 461 ArgGluAsnLysAspGluArgTyrSerIleLysAspLeuLeuAsnHisAlaPhePheGln 480
 Db 1381 CGACAAAACAAAGATGAAAGATATTCATCAAGACCTTTTGAACCATGCTCTCCAA 1440
 QY 481 GluGluThrGlyValArgValGluLeuAlaGluGluAspAspGlyGluLysIleAlaIle 500
 Db 1441 GAGAAACAGAGATACGGGTGAAATTAGCAGAGAAATGATGAGAAATAATAGCCATA 1500
 QY 501 LysLeuTyrPheLeuArgIleGluAspIleLysLysLeuLysGlyLysTyrLysAspAsnGlu 520
 Db 1501 AAATTTATGCTACGATATTGAAGATATTAAAGAAATTAAGGAAATATCAAAAGATATGAA 1560
 QY 521 AlaIleGluPheSerPheAspLeuGluArgAspValProGluAspValAlaGlnGluMet 540
 Db 1561 GCTATTAGTTTCTTTTGTATTGATAGAGAGATGCTCCAGAAAGATGTTGCAACAAAGATG 1620
 QY 541 ValGluSerGlyTyrValCysGluGlyAspHisLysThrMetAlaLysAlaIleLysAsp 560
 Db 1621 GTTAGAGTCTGGATGTCTGTGAAAGTGTATACAAAGCATGGCTAAAGCTATCAAGAC 1680
 QY 561 ArgValSerLeuIleLysArgGlyArgGluGlnArgGlnLeuValArgGluGluGln 580
 Db 1681 AGAGTATCATTAATTAGAGAGAAACGAGACAGCGGAGTTGGTGGAGAGACAAAGA 1740
 QY 581 LysLysLysGlnGluGluSerSerLeuLysGlnValGluGlnSerSerAlaSerGln 600
 Db 1741 AAAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1800
 QY 601 ThrGlyIleLysGlnLeuProSerAlaSerThrGlyIleProThrAlaSerThrThrSer 620
 Db 1801 AAGAGATCAAGAGCTCCCTCTGCTAGACACCGCATACCTACTGCTTCAACCACTCA 1860
 QY 621 AlaSerValSerThrGluValGluProGluGluProGluAlaAspGlnHisGlnGlnLeu 640
 Db 1861 GCTTCAGTTTCTACACAAGTAGAAGCTGAAGAACTGAGGCAAGTCAACATCAACACTA 1920
 QY 641 GlnTyrGlnGlnProSerIleSerValLeuSerAspGlyThrValAspSerGlyGlnGly 660
 Db 1921 CAGTACAGAGACCACTATATCTGTATCTGATGGAGCGTGTGACAGTGTGACAGGGA 1980
 QY 661 SerSerValPheThrGluSerArgValSerSerGlnGlnThrValSerTyrGlySerGln 680
 Db 1981 TCTCTGTCTTCAACGAATCTCAGTAGACAGCCACAGACAGTTTCAATAGGTTCCAA 2040
 QY 681 HisGluGlnAlaHisSerThrGlyThrValProGlyHisIleProSerThrValGlnAla 700
 Db 2041 CATGAACAGGCACATCTACAGAGACAGTCCAGGGGCAATACCTTCACTGTCACAGCA 2100
 QY 701 GlnSerGluProHisGlyValTyrProProSerSerValAlaGlnGlyGlnSerGlnGly 720
 Db 2101 CAGTCTCAGCCCATGGGTATATCAACCTCAAGGTGACACAGGGGAGACCGAGCGGT 2160
 QY 721 GlnProSerSerSerSerLeuThrGlyValSerSerSerGlnProIleGlnHisProGln 740
 Db 2161 CAGGCATCTCTCAAGTACCTTAACAGGGGTTCTCTCTCCAAACCATCAACATCTCTCAG 2220

QY 741 GlnGlnGlnGlyIleGlnGlnThrAlaProProGlnGlnIleThrValGlnTyrSerLeuSer 760
 Db 2221 CAGGACAGGGAATACAGACAGAGCCCTCTCAACAGACAGTGCATATTCACCTTCA 2280
 QY 761 GlnThrSerThrSerSerGluAlaThrThrAlaGlnProValSerGlnProGlnAlaPro 780
 Db 2281 CAGACATCAACCTCAAGTGAAGGCGCATCTGACAGCCAGTGAAGTCAAGCTCAAGCTCA 2340
 QY 781 GlnValLeuProGlnValSerAlaGlyLysGln----- 791
 Db 2341 CAACTTGGCTCAAGTATCAGCTGSAAAAACAGCTTCCAGTTTCCAGCAGTACCACT 2400
 QY 791 ----- 791
 Db 2401 ATCAAGGCGCACTCAGATCCAGTTCGACACACAAACCTCGTTGTTCCAGTCACTCT 2460
 QY 791 ----- 791
 Db 2461 GGTGCTCATTTCTTCCAGTGGAGACAGCGCTCCCTACTCCCTGCTCCCTCAATACCT 2520
 QY 791 ----- 791
 Db 2521 GTCTCTAGATTCCTCATATCAACTCTCATGTGTCTACGCTCAGACAGTTCCTCATCC 2580
 QY 791 ----- 791
 Db 2581 CTTCCTCATCAATAGGACGCTGGATTACTCAGCTCTGCTCAGCTTGGCTTCACTGCT 2640
 QY 791 ----- 791
 Db 2641 ACAACAGCTCGATCCCGGGGTATCAACTGTGTCTCTAGTACGTTCCAAACCTCTGTG 2700
 QY 791 ----- 791
 Db 2701 CAGCTGTGACTCAGCTGCCAAGTCAAGTTCACCCACAGCTCTCAACAGCAGATTCAAG 2760
 QY 791 ----- 791
 Db 2761 TCATGGGAATACCACTAAGTAACTTTGACAAAGCTGTGAGTTCCACTTCTCTGAGAT 2820
 QY 791 ----- 791
 Db 2821 GTTCTGTACAGAGGCTTCCCACTCGACTGCAACAGTACCAAGTACCAAGAGATTCAATATT 2880
 QY 791 ----- 791
 Db 2881 GCTCCCTCTTCAACAGTGGCTTGTGTGATCATTCATCAAGTCCATCCCTCCCATG 2940
 QY 791 ----- 791
 Db 2941 CCGACAGAACTACTGGCTACCTGGGTACTTTCACAGTGTGACAGCTTATGTGAA 3000
 QY 791 ----- 791
 Db 3001 TCAATCTTTAGTTCCTATGGGTGTGTGAGAGACAGGTTCAAGTCCAGCCAGCA 3060
 QY 792 ----- 792
 Db 3061 GGGAGTTTACAGCAAGCCCCCACTACATCTCCAGCAAGCAAGTTTGGAGAGTACTCAG 3120
 QY 795 GlyValSerGlnValAlaProAlaGluProValAlaValAlaGlnProGlnAlaThrGln 814
 Db 3121 GAGGTCTCAGGTTCTCTCTGACAGCCAGTTCGATGACAGACAGCCCAAGCTTACCCAG 3180
 QY 815 ProThrThrLeuAlaSerSerValAspSerAlaHisSerAspValAlaSerGlyMetSer 834
 Db 3181 CCAGCCACTTGGCTTCTCTGTAGACAGTGCACATTCAGATGTGCTTCAAGTATGAGT 3240
 QY 835 AspGlyAsnGluAsnValProSerSerSerArgHisGlnGlyArgThrThrLysArg 854
 Db 3241 GATGGCAATGAGAACTCCATCTTCCAGTGGAGAGCATGAAAGAACTCAACAAACGG 3300
 QY 855 HisTyrArgLysSerValArgSerArgSerArgHisGlnLysThrSerArgProLysLeu 874


```
Db 3301 CATTACCGAAATCTGTAGAGAGCGCTTCGACATGAAAAAACTTCACCCCAAAATTA 3360
Qy 875 Arg11IleuAsnValSerAsnLysGlyAspArgValValGluCysGlnLeuGluThrHis 894
Db 3361 AGAATTTTGAATGTTTCAAAATTAAGAGACCGAGTAGTAAATGTCATTAATAGACTCAT 3420
Qy 895 AsnArgLysMetValThrPheLysPheAspLeuAspGlyAspAsnProGluGluIleAla 914
Db 3421 AATAGGAAATGGTTACATTCATTAATTTGACCTAGATGTGACAAACCCCGAGAGATAGCA 3480
Qy 915 ThrIleMetValAsnAsnAspPheIleLeuAlaIleGluArgGluSerPheValAspGln 934
Db 3481 ACAATTGTGTGAACAATGACTTATCTAGCAATAGAGAGAGAGTGTGTTGTGATCAAA 3540
Qy 935 ValArgGluIleIleGluLysAlaAspGluMetLeuSerGluAspValSerValGluPro 954
Db 3541 GTGCGAATAATTTGTAAGAAAGCTGATGAATGCTCGTAGAGATGTCAGTGTGAAACCA 3600
Qy 955 GluGlyAspGlnGlyLeuGlnSerLeuGlnGlyLysAspAspTyrGlyPheSerGlySer 974
Db 3601 GAGGATGATCAGGAGATTGAGAGAGCTACAGAAAGAGATGACTATGCTTTTCAGGTTCT 3660
Qy 975 GluLysLeuGluGlyGluPheLysGlnProIleProAlaSerSerMetProGlnGlnIle 994
Db 3661 CAGAAATTTGGAGAGAGAGTCAAAACAACCAATTCCTGCTTCATGCCACAGCAAAATA 3720
Qy 995 GlyIleProHisSerSerLeuThrGlnValValHisSerAlaGlyArgGlyPheIleVal 1014
Db 3721 GGGATTCCTACCGAGTTTCTTAACCTCAAGTTGTTCTTCGCGAAGGCGGTTTATAGTG 3780
Qy 1015 SerProValProGluSerArgLeuArgGluSerLysValPheProSerGluIleThrAsp 1034
Db 3781 AGTCCTGTGCGAGAAAGCCGATTAACGAAATCAAAAGTTTCCCGAGTAAATACGAT 3840
Qy 1035 ThrValAlaAlaSerThrAlaGlnSerProGlyMetAsnLeuSerHisSerAlaSerSer 1054
Db 3841 ACAGTTGCTCCTCTACAGCTCAGAGCCCTGGAATGAACTGTCTCATCTGCATCATCC 3900
Qy 1055 LeuSerLeuGlnGlnAlaPheSerGluLeuArgAlaGlnMetThrGluGluProAsn 1074
Db 3901 CTTAGTCTACACAGGCGCTTTCTGAACTTAGAGAGTCCCAATAGACAGAAAGGCCAAC 3960
Qy 1075 ThrAlaProAsnProAsnPheSerHisThrGlyProThrPheProValValProPheLeu 1094
Db 3961 ACAGCACTCCAAACTTACTGATACAGAGCAACAACTTCCAGTAGTAACTCTCTTCTTA 4020
Qy 1095 SerSerIleAlaGlyValProThrThrAlaAlaAlaThrAlaProValProAlaThrSer 1114
Db 4021 AGTAGCATTTGCTGAGTCCCAACACAGCAGCAGCCACACACAGTCCCTGCACAAGC 4080
Qy 1115 SerProProAsnAspIleSerThrSerValIleGlnSerGluValThrValProThrGlu 1134
Db 4081 AGCCCTCTTAATGACATTTCCACATCGAATTCAGTGAAGTTAAAGGCCACCGAA 4140
Qy 1135 GluGlyIleAlaGlyValAlaThrSerThrGlyValValThrSerGlyLysLeuProIle 1154
Db 4141 GAGGGAGTTGCTGAGTGGCCACACAGCAGAGTGTGTAATCTCAGGTGCTCCCATTA 4200
Qy 1155 ProProValSerGluSerProValLeuSerSerValValSerSerIleThrIleProAla 1174
Db 4201 CCACCTGTGCTGATACACAGTACTTCCAGCTAGTTCATCAAGTATCAACAATCTCGCA 4260
Qy 1175 ValValSerIleSerThrThrSerProSerLeuGlnValProThrSerThrSerGluIle 1194
Db 4261 GTTGTCTCAATATCTACTACATCCCGTCACTTCAATCCCAATCCCAATCGAATC 4320
Qy 1195 ValValSerSerThrAlaLeuTyrProSerValThrValSerAlaThrSerAlaSerAla 1214
Db 4321 GTTGTCTCTAGTACAGACGTGATCCTTAGTAAATTTACAGAACTTACAGCTCTCGCA 4380
Qy 1215 GlyLysSerThrAlaThrProGlyProLysProProAlaValValSerGlnGlnAlaAla 1234
Db 4381 GGGGCGAGTACTGCTAACCCCGAGTCTTAAGCTTCAGCTGTAGTATCTCAGCAGCAGCA 4440
Qy 1235 GlySerThrThrValGlyAlaThrLeuThrSerValSerThrThrThrSerPheProSer 1254
Db 4441 GGCAGACACTACTGTGGAGCGCATTAACATCACTGTTTCTACCAACCACTTATCCCAAGC 4500
Qy 1255 ThrAlaSerGlnLeuSerIleGlnLeuSerSerSerThrThrProThrLeuAlaGlu 1274
Db 4501 ACAGCTTACAGCTGTGCATCTAGCTTAGCAGCAGTAACTTCACTCCTACTTAAAGTGAA 4560
Qy 1275 ThrValValValSerAlaHisSerLeuAspLysThrSerHisSerSerThrThrGlyLeu 1294
Db 4561 ACCGTGTGTGTGACCCACACTCCTACATGATTAAGACATCTCATGCACTACACTGATTC 4620
Qy 1295 AlaPheSerLeuSerAlaProSerSerSerSerSerProGlyAlaGlyValSerSerTyr 1314
Db 4621 GCTTCTCCTCTCTGACCATCTTCTCTCTCTCTGACACAGAGAGTGTCTAGTAT 4680
Qy 1315 IleSerGlnProGlyLysLeuHisProLeuValIleProSerValIleAlaSerThrPro 1334
Db 4681 ATTTCTCAGCTGTGGGGCTGCATCTTGTGTCATTCATCAATCAATAGCTTCTACTCT 4740
Qy 1335 IleLeuProGlnAlaAlaGlyProThrSerThrProLeuLeuProGlnValProSerIle 1354
Db 4741 ATTTCTCCCGACACAGCAGCAGCTTCTTACACTTATATCCCAAGTACTAGTATC 4800
Qy 1355 ProProLeuValGlnProValAlaAsnValProAlaValGlnGlnThrLeuIleHisSer 1374
Db 4801 CCACCTTGTGTACAGCTGTGTCCTGATGCTCTGCTGACAGACACATTAATTCATAGT 4860
Qy 1375 GlnProGlnProAlaLeuLeuProAsnGlnProHisThrHisCysProGluValAspSer 1394
Db 4861 CAGCTCAACCACTTGTGCTTCCCAACAGCCCAATACATCATTTGCTCGTAAGTATTC 4920
Qy 1395 AspThrGlnProLysAlaProGlyIleAspAspIleLysThrLeuGluGluLysArg 1414
Db 4921 GATACACAAACCAACCTCTGGAATGATGAATGAAGCTGTAGAGAAAGCTGCGG 4980
Qy 1415 SerLeuPheSerGlnHisSerSerSerGlyAlaGlnHisAlaSerValSerLeuGluThr 1434
Db 4981 TCTCTGTTAGTAAACACAGCTCATCTGAGCTCAGCATCTCTGTCTCATCTGAGACC 5040
Qy 1435 SerLeuValIleGlnSerThrValThrProGlyIleProThrThrAlaValAlaProSer 1454
Db 5041 TCACATGATCAAGAGCACTGTACACACAGGATCCCACTACTGCTGTGACCAAGC 5100
Qy 1455 LysLeuLeuThrSerThrThrSerThrCysLeuProProThrAsnLeuProLeuGlyThr 1474
Db 5101 AAACCTCTGACTTCTACACAGTACTGTGCTTACACCAACCAATTTACCACTAGGACA 5160
Qy 1475 ValAlaLeuProValThrProValValThrProGlyGlnValSerThrProValSerThr 1494
Db 5161 GTTGTCTTGCAGTTACACAGTGTGACACTGTGGCAAGTTTCTACCCAGTACAGCAT 5220
Qy 1495 ThrThrSerGlyValLysProGlyThrAlaProSerLysProProLeuThrLysAlaPro 1514
Db 5221 ACTACATCAGGAGTGAAGTGAAGTGAAGTGCCTCCACAGCCACTTAATTAAGGCTCG 5280
Qy 1515 ValLeuProValAlaGlyThrGluLeuProAlaGlyThrLeuProSerGluGlnLeuProPro 1534
Db 5281 GTGCTGCAGTGTGATCTGAATCTTCCAGCAGGTACTTACCCAGCAGCAGCTGCCACT 5340
Qy 1535 PheProGlyProSerLeuThrGlnSerGlnGlnProLeuGlnLysLeuAspAlaGlnLeu 1554
Db 5341 TTTCAGAGACTTCTCTTAACCACTCCAGCAACCTTGAAGATCTTGAATGCTCAATTC 5400
Qy 1555 ArgArgThrLeuSerProGluLysThrValThrSerAlaValGlyProValSerMet 1574
Db 5401 AGAAGAACTTATGTCAGAGATTATACAGTACTTCTGCGTGGTCTCTGTGTCATG 5460
Qy 1575 AlaAlaProThrAlaIleThrGluAlaGlyThrGlnProGlnLysGlyValSerGlnVal 1594
Db 5461 GCGGCTCCAAACCAATCAACAGAGCAAGAAACAGCTCAGAGAGGTGTTCTCAAGTC 5520
```

QY 1595 LysGluGlyProValLeuAlaThrSerSerGlyValGlyValPheLysMetGlyArgPhe 1614
DB 5521 AAGAGAGCCCTGCTCAGCACTAGTTCAGAGAGCTGGTGTGTTTAAAGAGGCGATTT 5580
QY 1615 GluValSerValAlaAlaAspGlyValGlnLysGlnLysAsnLysSerGluAspAla 1634
DB 5581 CAGGTTCTGTGTCAGAGAGCGGTGCCAGAAAGAGGTAAATTAAGTCAGAGATGCA 5640
QY 1635 LysSerValHisPheGlnSerSerThrSerGlnSerValLeuSerSerSerPro 1654
DB 5641 AAGCTGTTCATTGTAATCCAGACCTCAGAGCTCAGGCTCATCAAGTAGTAGTCCA 5700
QY 1655 GluSerThrLeuValLysProGlnLysProAsnGlyLysThrLysPheGlyLysSerAsp 1674
DB 5701 GAGAGTCTTGTGGAAACCAAGGCCAATAGCAATACCTCCTGGTATCTCTTCAGAT 5760
QY 1675 ValProGlnSerAlaHisLysThrThrAlaSerGlnAlaLysSerAspThrGlyGlnPro 1694
DB 5761 GTGCGAAGAGTGCCCAAAACTAGCTCCTCAGAGCAAAAGTCAGACACTGGGCAAGCT 5820
QY 1695 ThrLysValGlyArgPheGlnValThrThrAlaAsnLysValGlyArgPheSerVal 1714
DB 5821 ACCAAGGTGACGCTTTTCAGGTGACAACTACAGCAAAAGAGGTGCTGCTCTGTA 5880
QY 1715 SerLysThrGlnAspLysLysThrAspThrLysLysGlnGlyProValAlaSerProPro 1734
DB 5881 TCAAAACCTAGAGCAAGATCACTGACACAAAGAAAGAACAGCACTGGCATCTCTCCT 5940
QY 1735 PheMetAspLeuGlnGlnAlaValLeuProAlaValLysProLysGlnLysProGln 1754
DB 5941 TTTATGATTTGGAAACAAGCTGTTCTTCTGCTGTGATACCAAGAAAGAAAGCACTGAA 6000
QY 1755 LeuSerGlnProSerHisLeuAsnGlyProSerSerAspProGlnAlaAlaPheLeuSer 1774
DB 6001 CTGCAAGAGCTTACACTTAATGGAGGCGCTTCTTAACCCCGAGGCGCTTTTAAAGT 6060
QY 1775 ArgAspValAspAspGlySerGlySerProHisSerProHisGlnLeuSerSerLysSer 1794
DB 6061 AGGATGTGATGATGTTCCGATGATCCACACTGCCCACTGACCTGACCTCAAGAGC 6120
QY 1795 LeuProSerGlnAsnLeuSerGlnSerLeuSerAsnSerPheAsnSerLysMetSer 1814
DB 6121 CTTCCTAGCCAGAAATCTAAGTCAAAAGCCTTAGTAATCATTTAACTCCTCTTACATAGT 6180
QY 1815 SerAspAsnGlnSerAspLysLysLysAspGlnAspLysLeuLysLeuAlaGlnLeuArg 1834
DB 6181 AGCGACATAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6240
QY 1835 AspLysHisLeuLysGlnLysGlnAspLysSerArgGlnLysHisGlnLysGlnLys 1854
DB 6241 GATTAATATCTCAAAAGATGTCAGAGCTCAGAGCTGCGCAAGAGCTGAAATTTGATCT 6300
QY 1855 LeuLysThrLysLeuGlyLysValProProAlaValLysLeuProAlaAlaProLeu 1874
DB 6301 TTTGATTAACCAACTGGGCAAGGTGCCCTGCTGTATTTATCCCCAGAGCTGCCCTT 6360
QY 1875 SerGlyArgArgArgArgProThrLysSerLysGlySerLysSerSerArgSerSerSer 1894
DB 6361 TCGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6420
QY 1895 LeuGlyAsnLysSerProGlnLeuSerGlyAsnLeuSerGlyGlnSerAlaAlaSerVal 1914
DB 6421 TTGGGGAATAAAGCCCCAGCTTTCAGTATCACTGCTGCTGAGAGTCAAGCTTCAAGTC 6480
QY 1915 LeuHisProGlnGlnThrLeuHisProProGlyAsnLysProGlnSerGlyGlnAsnGln 1934
DB 6481 TTTCACCCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6540
QY 1935 LeuLeuGlnProLeuLysProSerProSerSerAspAsnLeuLysSerAlaPheThrSer 1954
DB 6541 CTGTATCAAGCCCTTAAAGCATCTCCTTCAAGTACAACTTATTCAGCTTCAACAGT 6600

QY 1955 AspGlyValAlaLysSerValProSerLeuSerAlaProGlyGlnGlyThrSerSerThrAsn 1974
DB 6601 GATGATCCCTTTTCAGTACCAAGCTTTCTGCTCAGGTCAAGAGACAGAGACCAAG 6660
QY 1975 ThrValGlyValAlaThrValAsnSerGlnAlaAlaGlnAlaGlnProProAlaMetThrSer 1994
DB 6661 ACTGTTGGGCAACAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6720
QY 1995 SerArgLysGlyThrPheThrAspLysPheHisLysLeuValAspAsnThrPalaArgAsp 2014
DB 6721 AGCAGAGAGGAGCAATTCACAGATGACTTCCACAAAGTTGTAAGCAATTGGGCGCGAGAT 6780
QY 2015 AlMetAsnLeuSerGlyArgArgGlySerLysGlyHisMetAsnArgGlnGlyProGly 2034
DB 6781 GCCATGATCTTCCAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6840
QY 2035 MetAlaArgLysPheSerAlaProGlyGlnLeuCysLysSerMetThrSerAsnLeuGly 2054
DB 6841 ATGGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6900
QY 2055 GlySerAlaProLysSerAlaAlaAlaSerAlaThrSerLeuGlyHisPheThrLysSerMet 2074
DB 6901 GGCTGCGCCCATCTCTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6960
QY 2075 CysProProGlnGlnArgGlyPheProAlaThrProPheGlyAlaGlnLysSerGlyThr 2094
DB 6961 TGCCCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7020
QY 2095 GlyGlyProAlaProGlnProLeuGlyGlnPheGlnProValGlyThrAlaSerLeuGln 2114
DB 7021 GTGCGCCAGACACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7080
QY 2115 AsnPheAsnLysSerLeuGlnLysSerLysSerAsnProProGlySerAsnLeuArg 2134
DB 7081 AATTTCAACATTCAGCAATTTGAGAAATCCATGAGCAACCCCGAGGCTTCAACTGCGG 7140
QY 2135 ThrThr 2136
DB 7141 ACCACT 7146

RESULT 3
US-10-491-467-54
Sequence 54, Application US/10491467
GENERAL INFORMATION:
APPLICANT: INCYTE CORPORATION; BANDMAN, Olga;
APPLICANT: BAUGHN, Mariah R.; BECHT, Shanya D.;
APPLICANT: BOROWSKY, Mark L.; DUGGAN, Brendan M.;
APPLICANT: EMERLING, Brooke M.; FORSYTHE, Ian J.;
APPLICANT: GANDHI, Ameena R.; GORVAD, Ann E.;
APPLICANT: GRIFPIN, Jennifer A.; GURURAJAN, Rajagopal;
APPLICANT: HAFALIA, April J.A.; KHAN, Parrish A.;
APPLICANT: LAL, Preeti G.; LEE, Ernestine A.;
APPLICANT: LEE, Soo Yeun; LINDQUIST, Erika A.;
APPLICANT: LU, Dying Anna M.; LU, Yan;
APPLICANT: MARQUIS, Joseph P.; NGUYEN, Daniel B.;
APPLICANT: MARVIZU, Chandra S.; RAMKUMAR, Jayalaxmi;
APPLICANT: RECIPON, Shirley A.; RICHARDSON, Thomas W.;
APPLICANT: SWARNAKAR, Anita; TANG, Y. Tom;
APPLICANT: THORNTON, Michael B.; TRAN, Uyen K.;
APPLICANT: CHAWLA, Narinder K.; WARREN, Bridget A.;
APPLICANT: YANG, Junming; YAO, Monique G.;
APPLICANT: YUE, Henry; ZEBARUDIAN, Yeganeh
TITLE OF INVENTION: KINASES AND PHOSPHATASES
FILE REFERENCE: PF-1244 USN
CURRENT APPLICATION NUMBER: US/10/491,467
CURRENT FILING DATE: 2004-03-31
PRIOR APPLICATION NUMBER: PCT/US02/33723
PRIOR FILING DATE: 2002-10-17
PRIOR APPLICATION NUMBER: US 60/345,474
PRIOR FILING DATE: 2001-10-19
PRIOR APPLICATION NUMBER: US 60/343,910
PRIOR FILING DATE: 2001-11-02
PRIOR APPLICATION NUMBER: US 60/333,098

QY ThrGlyLeuValGlnProSerAlaSerThrGlyLeuProThrAlaSerThrThrSer 620
Db AGAGGATCAAGACGCTCCTTCTGCTAGCACCGCATCTACTGCTTCTTACCACTTCA 2220
QY AlaSerValSerThrGlnValGlnProGlnGluProGlnAlaAspGlnHisGlnGlnLeu 640
Db GCTTCAGTTTCTCAACAAGTAGAACCTGAAGAACCTGAGCAGATCAACATCAACAATA 2280
QY GlnTyrGlnGlnProSerIleSerValLeuSerAspGlyThrValAspSerGlyGlnGly 660
Db CAGTACAGAACCCAGTATATCTGTGTATCTGATGGACCGGTTCACATGTGTCAAGGA 2340
QY SerSerValPheThrGlnSerArgValSerSerGlnGlnThrValSerTyrGlySerGln 680
Db TCTCTGTCTTCAACAGATCTCCAGTGACAGCCACACAGCACTTTCATGTGTCCCA 2400
QY HisGlnGlnAlaHisSerThrGlyThrValProGlyHisIleProSerThrValGlnAla 700
Db CATGAACAGGCATCTTCAAGCAGCAGTCCAGGGCATATCTTCTACTGTCCAAGCA 2460
QY GlnSerGlnProHisGlyValTyrProProSerSerValAlaGlnGlnGlnSerGlnGly 720
Db CAGTTCAGGCCCATGGGTATATCCACCTCAAGTGTG----- 2499
QY GlnProSerSerSerSerLeuThrGlyValSerSerSerGlnProIleGlnHisProGln 740
Db ----- 2499
QY GlnGlnGlnGlnGlyIleGlnGlnThrAlaProProGlnGlnThrValGlnTyrSerLeuSer 760
Db -----CACAGAGGATATACAGCAGACAGCCCTCTCAACAGACAGTGCAGTATCACTTCA 2556
QY GlnThrSerThrSerSerGlnAlaIleThrThrAlaGlnProValSerGlnProGlnAlaPro 780
Db CAGACATCAACCTCCAGTGAGGCCCATCTACTGACAGCAGTGTGATCACTCAAGCTCCA 2616
QY GlnValLeuProGlnValSerAlaGlyIleGlnSerThrGlnGlnValSerGlnValAla 800
Db CAAGTCTTGCTCAAGATACGTGAGAAAACAGATCTCAGGAGTCTCTCAGGTTGCT 2676
QY ProAlaGlnProValAlaValAlaGlnProGlnAlaThrGlnProThrThrLeuAlaSer 820
Db CTGACAGAGCCAGTGTGAGTAGCACAGCCCAAGCTTCCAGCCGACCACTTGGCTTCC 2736
QY SerValAspSerAlaHisSerAspValAlaSerGlyMetSerAspGlyAsnGlnVal 840
Db TCTGTAGACAGTGCATTCAGATGTGTCTTCAAGTATGAGTATGGCAATGGAACGTC 2796
QY ProSerSerSerGlyArgHisGlnGlyArgThrThrLysArgHisTyrArgLysSerVal 860
Db CCATCTTCCAGTGAAGGATGAAGGAAGAACCAAAACGGCATTTACCAAAATCTGTA 2856
QY ArgSerArgSerArgHisGlnLysThrSerArgProLysLeuArgIleLeuAsnValSer 880
Db AGAGAGTGCCTCTGACATGAATAAACTTCAAGCCCAAAATTAAAGATTGTAATGTTTCA 2916
QY AsnLysGlyAspArgValValGlnCysGlnLeuGlnThrHisAsnArgLysMetValThr 900
Db AATTAAGGAAGACCGAGTAGAATGTCAATTAGAGCTCATATATAGAAAAAGGTTTACA 2976
QY PheLysPheAspLeuAspGlyAspAsnProGlnGlnIleAlaThrIleMetValAsnAsn 920
Db TTTCAATTTTCACTAGATGTGTGACACCCCGAGAGATGACCAACATTAATGTGAAACAT 3036
QY AspPheIleLeuAlaIleGlnArgGlnSerPheValAspGlnValArgGlnIleIleGln 940
Db GACTTATTTCTAGAAATAGAGAGAGAGTCTGTTGTGATCAAGTGCAGAAATTAATGAA 3096
QY LysAlaAspGlnMetLeuSerGlyAspValSerValGlnProGlnGlnLysAspGlnGlyLeu 960
Db AAGGTCATGAAATGCTCAGTGAAGATGTCAAGTGTGAACCAAGGGGTGATTCAGGGATTG 3156
QY GlnSerLeuGlnGlnLysAspAspTyrGlyPheSerGlySerGlnLysLeuGlnGlnGlu 980

Db 3157 GAGAGCTCAAGAGAAAGATGACTATGCTTTTCAAGTTTCTCAAGAAATTGGAAGAGAG 3216
QY PheLysGlnProIleProAlaSerSerMetProGlnGlnIleGlyIleProThrSerSer 1000
Db TTTCAACCAACCAATTCCTCGTCTTCCATGCCACACCAATATAGCATTCCTTACAGTTCT 3276
QY LeuThrGlnValValHisSerAlaGlyArgAspPheIleValSerProValProGlnSer 1020
Db TTAATCAAGTTGTATATCTGGGAGAGCGGTATATAGTGTCTGTGCGCAAGAAC 3336
QY ArgLeuArgGlnSerLysValPheProSerGlnIleThrAspThrValAlaAlaSerThr 1040
Db CCAATTACAGAAATCAADAAGTTTCCCAAGTAATAACAGATACAGTTGTGCTCTACAC 3396
QY AlaGlnSerProGlyMetLeuLeuSerHisSerAlaSerSerLeuSerLeuGlnGlnAla 1060
Db GCTCAGAGCCCTGGAAATGAATGTCTCACTGCAATCATCCCTTGTCTACAAACAGGCC 3456
QY PheSerGlnLeuArgArgAlaGlnMetThrGlnGlyProAsnThrAlaProProAsnPhe 1080
Db TTTTCTGAATTTAGAGTGTGCCAAATGACAGAAAGAACCAATACAGCACTTCCAAACTTT 3516
QY SerHisThrGlyProThrPheProValValProProPheLeuSerIleAlaGlyVal 1100
Db AGCATACAGAACCAACATTTCCAGTAGTACCTCTTTTATAGTACATGCTGGAGTCC 3576
QY ProThrThrAlaAlaAlaThrAlaProValProAlaThrSerSerProProAsnAspIle 1120
Db CCAACCAAGCAGACAGCCACAGCAGCAGTCCGCAACAAAGCCCTCCTAATATGACAT 3576
QY SerThrSerValIleGlnSerGlnValThrValProThrGlnGlnGlyIleAlaGlyVal 1140
Db TCCATCATAGTATTAAGTCTGAGTTACAGTCCCACTGACAGAGGGATTCGTGAAGTT 3596
QY AlaThrSerThrGlyValValThrSerGlyGlyLeuProIleProProValSerGlnSer 1160
Db GCCACAGACAGAGTGTGTAATCTCAGGTGTCTCCCATACCACTGCTGTAATCA 3756
QY ProValLeuSerSerValValSerSerIleThrIleProAlaValValSerIleSerThr 1180
Db CCAGTCTTCCAGCGTATGTTCAAGTATCAAAATCCAGCAATGTGTCTCAATATCTACT 3816
QY ThrSerProSerLeuGlnValProThrSerThrSerGlnIleValValSerSerThrAla 1200
Db ACATCCCGTCACTTAAGTCTCCCAATCCATCCATCGAATCGTGTCTTATGATACGCA 3876
QY LeuTyrProSerValThrValSerAlaThrSerAlaSerAlaGlySerThrAlaThr 1220
Db CTGTATCTTCAAGTAAAGTTTACAGCACTTCAAGCTTCCAGGGGACAGTACTGCTACC 3936
QY ProGlyProLysProProAlaValValSerGlnGlnAlaAlaIleLysSerThrThrValGly 1240
Db CCAGGTCTCAAGCTTCAGTGTAGTATCTCAGACAGGACAGCAGCACTACTGTGGGA 3996
QY AlaThrLeuThrSerValSerThrThrThrSerPheProSerThrAlaSerGlnLeuSer 1260
Db GCCACATTAACATCAAGTTTCTACCACTTCAATTCACAGCAGCTTACAGCTGTCC 4056
QY IleGlnLeuSerSerSerThrSerThrProThrLeuAlaGlnThrValValSerAla 1280
Db ATTCAGATTAGCAGCACTTCTACTCTCACTTGTAGCTAAACCGGTGATTAAGCGCA 4116
QY HisSerLeuAspLysThrHisSerHisSerSerThrThrGlyLeuAlaPheSerLeuSerAla 1300
Db CACTCACTATATAGACATCTCATAGACATGACAACTGATGTGCTTCTCTCTCTGCA 4176
QY ProSerSerSerSerSerProGlyValAlaGlyValSerSerTyrIleSerGlnProGlyGly 1320
Db CCATCTTCTCTCTCTCTCTCTGAGCAGAGAGTGTCTAGTTATTTCTACGCTGTGGGG 4236
QY LeuHisProLeuValIleProSerValIleAlaSerThrProIleLeuProGlnAlaAla 1340

Db 4237 CTGCATCTTTGGTCATTCATCATGATGATGCTTCTAATCTCTTATCTTCCCAAGCAGCA 4296
Qy 1341 GlyProThrSerThrProLeuLeuProGlnValProSerIleProProLeuValGlnPro 1360
Db 4297 GGACCTACTTCTACACCTTTATTAACCCCAAGTACCTTAGTATCCACCCCTTGTAACGCT 4356
Qy 1361 ValAlaAsnValProAlaValGlnGlnThrLeuIleHisSerGlnProGlnProAlaLeu 1380
Db 4357 GTTGCATATGCTGCTGTACAGACAGACACTAATTCATAGTCAGCTCAACCAAGCTTG 4416
Qy 1381 LeuProAsnGlnProHisThrHisGlySerProGlnValAspSerAspThrGlnProValAla 1400
Db 4417 CTTCACCAAGCCCACTACCTATCTGCTGAAGTATCTGATACCAACCAAGCT 4476
Qy 1401 ProGlyIleAspAspIleLeuThrLeuGlnGlnIleLeuArgSerLeuPheSerGlnHis 1420
Db 4477 CTTGGAATTCATACATTAAGACTCTAGAGAAAAGCTGCGGTCTCTGTTCAAGTGAACAC 4536
Qy 1421 SerSerSerGlyValGlnHisAlaSerValSerLeuGlnThrSerLeuValIleGlnSer 1440
Db 4537 AGCTCATCTGAGCTGAGCTGCTCTGCTCACTGAGACCTCACTAGTCATAGAGAGC 4596
Qy 1441 ThrValThrProGlyIleProThrThrAlaValAlaProSerIleLeuThrSerThr 1460
Db 4597 ACTGTACACCAAGCATCCCACTACCTGCTTGCAACCAAGCAAACTCTGACTTCTAAC 4656
Qy 1461 ThrSerThrCysLeuProProThrAsnLeuProLeuGlyThrValAlaLeuProValThr 1480
Db 4657 ACAAGTACTGCTTACCAACCAATTTCCACTAGGAACAGTTGCTTTGCCAGTTACA 4716
Qy 1481 ProValValThrProGlyGlnValSerThrProValSerThrThrThrSerGlyValHis 1500
Db 4717 CCAGTGTACACCTGGGCAAGTTTCTACCCAGTCAGCACTACATCAGAGAGTGA 4776
Qy 1501 ProGlyThrAlaProSerIleProProLeuThrIleAlaProValLeuProValGlyThr 1520
Db 4777 CCTGGAACTGCTCCCTCCAGCAACCTTACTAAGGCTCGGTGCTGCCAGTGGGACT 4836
Qy 1521 GlnLeuProAlaGlyThrLeuProSerGlnGlnLeuProProPheProGlyProSerLeu 1540
Db 4837 GAATCTTCACAGGATCTTACCCAGGAGAGCTGCACTTTTCCAGAACCTTCTCTA 4896
Qy 1541 ThrGlnSerGlnGlnProLeuGlnAspLeuAspAlaGlnLeuArgThrLeuSerPro 1560
Db 4897 ACCCAGTCCAGCACTCTAGAGGATCTTGATGCTCAATTGAGAGAAACACTTACGCCA 4956
Qy 1561 Gln***IleThrValThrSerAlaValAlaGlyProValSerMetAlaAlaProThrAlaIle 1580
Db 4957 GAGATTATCACAGTGACTTCTGCGGTGCTCTGTCTCCATGGCGGCTCCAACAGCATC 5016
Qy 1581 ThrGlnAlaGlyThrGlnProGlnIleValSerGlnValIleGlnGlyProValLeu 1600
Db 5017 ACGAAGCAGGAACACAGCTCCAGAGAGGAGTCTTAAGTCAAAAGAGGCCCTGTCTCA 5076
Qy 1601 AlaThrSerSerGlyAlaGlyValPheIleMetGlyArgPheGlnValSerValAlaAla 1620
Db 5077 GCAACTAGTTCAGAGAGCTGAGTGTTTTAAGATGGAGAGATTTCAGGTTCTGTGTGACGA 5136
Qy 1621 AspGlyValaGlnIleGlnGlyIleValAsnIleSerGlnAspAlaIleSerValIleIlePheGln 1640
Db 5137 GACGGTCCCAAGAAAGGGGTAAATAATAGTCAGAAATGCAAAAGTCTGTCTATTTTGAA 5196
Qy 1641 SerSerThrSerGlnSerSerValLeuSerSerSerSerProGlnSerThrLeuValHis 1660
Db 5197 TCCAGACCTTCAGAGTCTCAGTGTCTCAAGTATGTCACAGAGTACCTTGGTGAAA 5256
Qy 1661 ProGlnProAsnGlyIleThrIleProGlyIleSerSerAspValProGlnSerAlaHis 1680
Db 5257 CCAAGAGCAATGGCATTAACCATCCCTGATATCTTCAATGTCAGAGAGTGGCCAC 5316
Qy 1681 LysThrThrAlaSerGlnAlaIleValSerAspThrGlyGlnProThrIleValGlyArgPhe 1700
Db 5317 AAAACTATCTGCTTCAGGGCAAGTACAGACATGGGAGCCCTACCAAGGTTTGACGTTT 5376

Qy 1701 GlnValThrThrThrAlaAsnIleValGlyArgPheSerValSerIleThrGlnAspLys 1720
Db 5377 CAGGTACAACTACAGCAAAACAAAGGGTGGTCTGTCTGTATCAAAAACTGAGAGCAAG 5436
Qy 1721 IleThrAspThrIleValGlnGlyProValAlaSerProProPheMetAspLeuGln 1740
Db 5437 ATCAGTACCAAAAGAAAGAGGACAGTGGCATCTCTCTTTATGTGATTGGAAACAA 5496
Qy 1741 AlaValLeuProAlaValIleProIleValGlnGlyValProGlnLeuSerGlnProSerHis 1760
Db 5497 GCTGTCTTCTGCTGTGATACCAAGAAAGAAAGACCTGAATCTCAGAGCTTCAAT 5556
Qy 1761 LeuAsnGlyProSerSerAspProGlnAlaAlaPheLeuSerArgAspValAspAspGly 1780
Db 5557 CTAAATGGGCGGCTTCTGACCCCGAGGCGGCTTTTAAAGTAGGATGTGATGATGT 5616
Qy 1781 SerGlySerProHisSerProHisGlnLeuSerSerIleSerLeuProSerGlnAsnLeu 1800
Db 5617 TCCGATAGTCCACACTGCCCCATCAGCTGAGCTCAAGAGCTTCTAGCCAGAACTTA 5676
Qy 1801 SerGlnSerLeuSerSerPheAsnSerSerTyrMetSerSerAspAsnGlnSerAsp 1820
Db 5677 AGTCAAGCCTTAGTATTCATTAACTCTCTTACATAGATAGCAATGAGTACAT 5736
Qy 1821 IleGlnAspGlnAspLeuIleValLeuArgArgLeuArgAspLysHisLeuLysGln 1840
Db 5737 ATGGAAGATGAAGACTTAAGTTAGAGCTGCGACGACTAGATATAACATCTCAAGAG 5796
Qy 1841 IleGlnAspLeuGlnSerArgGlnIleValIleGlnSerLeuThrIleValGly 1860
Db 5797 ATTCAGAGACTGAGAGTGGCCAGAAAGCATGAATGAATCTTGTATACCAACTGGGC 5856
Qy 1861 LysValProProAlaValIleIleProProAlaAlaProLeuSerGlyArgArgArg 1880
Db 5857 AAGTCCCCCTGCTGTATATATCTCCCAAGCTGCTCCCTTACAGGAAACAGCA 5916
Qy 1881 ProThrLysSerIleValSerIleValSerSerArgSerSerSerLeuGlnAsnLysSerPro 1900
Db 5917 CCACTTAAGCAAGCAAGCAAGCAATCTAGTCAGAGAGTCTCTGGGGAATTAAGCCCC 5976
Qy 1901 GlnLeuSerGlyValAsnLeuSerGlyGlnSerAlaAlaSerValLeuHisProGlnGlnThr 1920
Db 5977 CAGCTTTCAGGTAACTGTGTGTGTCAGAGTCCAGTTCACTGTTGACCCCGAGAGAC 6036
Qy 1921 LeuHisProProGlyValIleProGlnSerGlyGlnAsnGlnLeuGlnProIleLys 1940
Db 6037 CTCACCTTCTGGCAACATCCAGAGTCCGGGCAAGTCAAGTGTACAGCCCCCTTAAG 6096
Qy 1941 ProSerProSerSerAspAsnLeuTyrSerAlaPheThrSerAspGlyAlaIleSerVal 1960
Db 6097 CCATCTCCCTCCAGTACAACTCTATTCAGCTTACCAAGTATGATGTCATTCAGTA 6156
Qy 1961 ProSerLeuSerAlaProGlyGlnGlyThrSerSerThrAsnThrValGlyAlaThrVal 1980
Db 6157 CCAAGCCTTCTGCTCCAGGTCAAGGAACAGAGCAACAAACCTGTTGGGCAACAGTG 6216
Qy 1981 AsnSerGlnAlaAlaGlnIleGlnProProAlaMetThrSerSerArgLysGlyThrPhe 2000
Db 6217 AACAGCAAGCCGCCCAAGCTCAGCTCTGCTCATACGTCACAGAGAAAGGCACTTC 6276
Qy 2001 ThrAspAspLeuHisIleValLeuValAspAsnTyrAlaArgAspAlaMetAsnLeuSerGly 2020
Db 6277 ACAGATGACTTCAGCAAGATTGTAGACAAATTTGGCCCGAGATCCCATGATCTCAAGGC 6336
Qy 2021 ArgArgGlySerIleValIleValMetAsnTyrGlnGlyProGlyMetAlaArgLysPheSer 2040
Db 6337 AGCAGAGGAAGCAAGGCAATGAATTAAGAGGCTCGGAATGCAAGAGAGTTCTCT 6396
Qy 2041 AlaProGlyGlnLeuValIleSerMetThrSerAsnLeuGlyValLysSerAlaProIleSer 2060
Db 6397 GCACTGGGCACTGTGATCTCCATGACCTCAACCTGGGTGCTGCTGCCCATCTCT 6456

Db 2396 GAGGAAACAGAGTACGGTGAATTAACGAAAGAGATGTGAGAGAAAAATAGCCATA 2455
 QY 501 LysLeuTripleuArg1IleGluAsp1IleValysLeuIysGlyIysIysAspAsnGlu 520
 Db 2456 AAATAATGGCTACCTATTAAGATATTAAGAAATTAAGGAAATATCAAGATATGA 2515
 QY 521 Ala1IleGluPheSerPheAspLeuGluArgAspValProGluAspValIleGluIleMet 540
 Db 2516 GCATATGAGTTTCTTTGATTTAGAGAGATGTCCAGAAAGATGTGGCAAGAAATG 2575
 QY 541 ValGluSerGlyIysValIysGluIysAspIleIysThrMetAlaIleValAsp 560
 Db 2576 GTAGAGTCGGTATGCTGTGAAGGTGATCAAGACCAATGGCTAAAGCTATCAAGAAC 2635
 QY 561 ArgValSerLeuIleValSerArgIysArgGluIleValIleGluIleValIleGluIle 580
 Db 2636 AGAGTATCATTAATTAAGAGAAACGAGACAGCGGAGTTGGTACGGAGAGAGAA 2695
 QY 581 LysIysIysGluIleGluIleGluIleGluIleGluIleValIleGluIleSerIleAspGln 600
 Db 2696 AAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2755
 QY 601 ThrGlyIleIysGluIleProSerIleAspThrGlyIleProThrIleAspThrIleSer 620
 Db 2756 ACAGGAATCAAGACGCTCCCTCTGTGAGACACCGGCACTACTGCTTACCACTCA 2815
 QY 621 AlaSerValSerThrGluValGluProGluIleGluIleValIleAspGlnIleGluIle 640
 Db 2816 GCTTCACTTTCTACACAGTATGAAGACCTGAAGAACCTGAGCAATCAACATCAACACTA 2875
 QY 641 GlnIysGluIleProSerIleSerValIleSerAspGlyIysValIleAspSerGlyIleGly 660
 Db 2876 CAGACACAGACACCACTATATCTGTATCTGATGGAGACGGTTGACAGTGTGACAGGA 2935
 QY 661 SerSerValPheThrArgIysArgValSerSerGluIleIleValIleSerIysGlySerGln 680
 Db 2936 TCCCTGTCTTCAVAGATCTCGAGTACGACGACCAAGACAGATTTATATATGTTCCAA 2995
 QY 681 HisGluIleAlaHisSerThrGlyIysThrValProGlyHisIleProSerThrValIleAla 700
 Db 2996 CAGACACAGACACCTTCAAGACGACAGTCCAGGACATATCTTACTGTCCAGCA 3055
 QY 701 GlnSerGlnProHisGlyIysIysProProSerSerValIleGluIleGlnIleGlnIleGly 720
 Db 3056 CAGCTCAGCCCGCATGGGTATATCCACCTCAAGTGTGGACAGGGGACAGAGCCAGGT 3115
 QY 721 GlnProSerSerSerSerSerIleThrGlyIleValSerSerSerGlnProIleGlnIleProGln 740
 Db 3116 CAGCCATCTCTCAAGTACTTAACAGGGGTTTCATCTTCCCAACCATTAACAATCCCTCAG 3175
 QY 741 GlnGlnGlnGlyIleGlnIleIleThrAlaProProGlnIleIleValIleGlnIysSerLeuSer 760
 Db 3176 CAGACACAGGAAATACAGACAGACAGCCCTCTCAACAGACAGTGCAGTATCTTCA 3235
 QY 761 GlnThrSerThrSerSerGluAlaIleThrThrAlaGlnProValSerGlnProGlnIleAlaPro 780
 Db 3236 CAGACATCAACCTCAGTGAAGCCACTACTGCAAGACAGAGTGTATCACTCAAGCTCCA 3295
 QY 781 GlnValIleupProGlnIleValSerIleAlaGlyIysGln----- 791
 Db 3296 CAAGTCTTGGCTCAAGTATCAGTGTGAAACAGCTTCCAGTTTCCAGCCAGTACCAACT 3355
 QY 791 ----- 791
 Db 3356 ATCCAGGCGAAGCTCAGATCCAGTTGCCAGACACAACTCCGTTGTTCCAGTCACTCT 3415
 QY 791 ----- 791
 Db 3416 GGTGCTCATTTCTTCCAGTGGAGACAGCGCTCCCTACTCCCTGTCTCCCTCAGTACCGCT 3475
 QY 791 ----- 791
 Db 3476 GTCTCTCATATTCATATCAACTCTCAGTGTCTACGAGCTCAGACAGAGTTTCTCATCC 3535

QY 791 ----- 791
 Db 3536 CTTCATATCAATGAGAGCTGGCATTTACTCAGCCTCTGTCTCAGTGTGCTCATCTGCT 3595
 QY 791 ----- 791
 Db 3596 ACAACAGCTGCATCCCGGGGTATCAACTGTGGTTCCTAGTACGTTCCAAACCTTCTG 3655
 QY 791 ----- 791
 Db 3656 CAGCTGTGACTAGCTGCCAAGTCAAGTTCAACCCACAGCTCTCTACACAGACAGTTTCAG 3715
 QY 791 ----- 791
 Db 3716 TCCATGGGAATACAGCTAACCTTGGACAAAGCTGTGAGGTTCCACTTCTCTGAGAT 3775
 QY 791 ----- 791
 Db 3776 GTTCTGTACAGGGCTTCCACCTCGACTGCCACACAGTACCCAGAGATTCAAATATT 3835
 QY 791 ----- 791
 Db 3836 GCTCCCTCTTCAACGTGCTTCTGTTGATCCATCTTAACAGTCTTATCCCTCCATG 3895
 QY 791 ----- 791
 Db 3896 CCGACAGAAATACGTGGCTACACCTGGGTACTTCCACAGTGGGACGCTTATGTGAA 3955
 QY 791 ----- 791
 Db 3956 TCAATCTTTTACTTCTATGAGGTGTGAGAGACAGGTCAAGTCCACGCCAGCA 4015
 QY 792 ----- SerThrGln 794
 Db 4016 GGGAGTTTACACAGGCCCCCACTACCTCTCCAGACACAGTTTGGAGAGTACTCAG 4075
 QY 795 GlyValSerGlnValAlaProAlaGluProValAlaValAlaGlnProGlnIleAlaThrGln 814
 Db 4076 GAGGTCTCTCAGGTGCTCTGTCAGAGCCAGTTGCACTACACAGMCCAAAGTACCCAG 4135
 QY 815 ProThrThrIleuAlaSerSerValAspSerAlaHisSerAspValAlaSerGlyMetSer 834
 Db 4136 CCGACCACTTGGCTCTCTCTGTAGACAGTGCACATTCATGTTGCTTCAAGTATAGT 4195
 QY 835 AspGlyAsnGluAsnValProSerSerSerGlyAlaHisGluGlyArgThrThrIleValArg 854
 Db 4196 GATGGCATAGAGACGTCCTCATCTTCCAGTGAAGGATGAGAGAGAACTTCAAAACGG 4255
 QY 855 HisIysArgIysSerValArgSerArgSerArgHisGluIysThrSerArgProIysLeu 874
 Db 4256 CATTACGGAATCTGTAAAGAGTCCGCTCGACATGAATAAACTTCAAGCCCAAAATTA 4315
 QY 875 ArgIleLeuAsnValSerAsnIysGlyAspArgValAlaGluIysGlnIleGluIleHis 894
 Db 4316 AGAATTTGATGTTTCAATAAAGAGACCGGTGTAGTAAGTCAATTAGAGACTCAT 4375
 QY 895 AsnArgIysMetValThrPheIysPheAspLeuAspGlyAspAsnProGluIleAla 914
 Db 4376 AATRGAAATATGTTTCAATTCAATTGACCTGATGTATGACACCCCGAGAGATATGCA 4435
 QY 915 ThrIleMetValAsnAsnAspPheIleLeuAlaIleGluIysGluIlePheValAspGln 934
 Db 4436 ACAATTATGTGAACATGACTTATCTGACAAATAGAGAGAGGTGTTGTGTGATCAA 4495
 QY 935 ValArgGluIleIleGluIysAlaAspGluMetLeuSerGluAspValSerValGluPro 954
 Db 4496 GTGCGAAGAAATATTATGAAAAAGCTGATGAATGTCTGATGATGTCAAGTGTGAACCA 4555
 QY 955 GluGlyAspGlnGlyIleGluIleGluIleGluIleGluIleGluIleGluIleGluIleGluIle 974
 Db 4556 GAGGATATCAGGGAATTTGAGAGTCTACAGAGAAAGATTAAGTATGCTTTTCAAGTTCT 4615

QY 975 GlnIysLeuGlnIyGluPheIysGlnProIleProAlaSerSerMetProGlnIle 994
Db 4516 CAGAAATGGAAGAGAGAGTCAAAACCAATTCCTGCTTCCTCATGCCACAGAAATA 4675
QY 995 GlyIleProThrSerSerLeuThrGlnValValHisSerAlaGlyArgArgPheIleVal 1014
Db 4676 GGCAATCTCTACAGATCTTTTAACCAAGTTGTTCAATCTCGGGAGAGCGGTTTATATGTG 4735
QY 1015 SerProValProGluSerArgLeuArgGluSerIysValPheProSerGluIleThrAsp 1034
Db 4736 AGTCTGTGCGCAAGAACCGCATTCAGAAATCAAAAGTTTCCCAAGTGAATATACAGAT 4795
QY 1035 ThrValAlaAlaSerThrAlaGlnSerProGlyMetAsnLeuSerHisSerAlaSerSer 1054
Db 4796 ACAATTGCTGCTCTACAGCTCAGAGCCCTGGAAATGAATCTGTCTCATCTCGCATATCC 4855
QY 1055 LeuSerLeuGlnGlnAlaPheSerGluLeuArgArgAlaGlnMetThrGluGlyProAsn 1074
Db 4856 CTTAGCTTACAAAGCGCTTTTCTGAACCTTAGAGCTGCCAAATGACAGAGAGACCCAAAN 4915
QY 1075 ThrAlaProProAsnPheSerHisThrGlyProThrPheProValValProProPheLeu 1094
Db 4916 ACAGCACTCCAACTTACATTCAGCATACAGACCAACATTTCCAGTAGTACCTCTTCTTA 4975
QY 1095 SerSerIleAlaGlyValProThrThrAlaAlaAlaThrAlaProValProAlaThrSer 1114
Db 4976 AGTAGATTTGGTGGAGTCCCAACACAGACAGACAGACAGACAGACAGACAGACAGAC 5035
QY 1115 SerProProAsnAlaPheSerThrSerValIleGlnSerGluValThrValProThrGlu 1134
Db 5036 AGCCCTCTTAATGACATTTCCACATCAGTAATTCAGCTGAGTTACAGTCCACACGAA 5095
QY 1135 GluGlyIleAlaGlyValAlaThrSerThrGlyValValThrSerGlyIleProIle 1154
Db 5096 GAGGGAGTTGCTGAGATGCCACACAGACAGAGTGTGTACTTACAGTGGCTTCCCATTA 5155
QY 1155 ProProValSerGluSerProValLeuSerSerValValSerSerIleThrIleProAla 1174
Db 5156 CCACCTGTGTGATACACACAGTACTTCCAGCTAGTTCAAGTATCACAAATACCTGCA 5215
QY 1175 ValValSerIleSerThrThrThrSerProSerLeuGlnValProThrSerThrSerGluIle 1194
Db 5216 GTTGTCTCAATCTACTTACATCCCGCTCACTCAACTCCCACTCCACATCTGCAATC 5275
QY 1195 ValValSerSerThrAlaLeuIleThrProSerValThrValSerAlaThrSerAlaSerAla 1214
Db 5276 GTTGTCTCTAGTACAGACCTGTATCTTCACTTCAAGTATTCAGCAACTTCAAGCTCTGCA 5335
QY 1215 GlyIleSerThrAlaThrProGlyProIleProProAlaValValSerGlnIleAla 1234
Db 5336 GGGGGGAGTACTGCTACCCAGAGCTCTAAGCTCCAGCTGATGATCTCAGAGGAGCA 5395
QY 1235 GlySerThrThrValGlyAlaThrLeuThrSerValSerThrThrThrSerPheProSer 1254
Db 5396 GGGAGCACTACTGGGAGCACATTAACATCAATTTCTACACACACTTCAATTCACAGC 5455
QY 1255 ThrAlaSerGlnLeuSerIleGlnLeuSerSerSerThrSerThrProThrLeuAlaGlu 1274
Db 5456 ACAGCTTCAAGCTGTGSCATTCACGCTTAGACAGCATCTTCTACTCTTACTTTAGCTGAA 5515
QY 1275 ThrValValValSerAlaHisSerLeuAspIleThrSerHisSerSerThrThrGlyLeu 1294
Db 5516 ACCGTGTAGTTCAGCAGACACTCACTAGATACATCTCATAGCAGTACCAACTGATTTG 5575
QY 1295 AlaPheSerLeuSerAlaProSerSerSerSerSerProGlyAlaGlyValSerSerTyr 1314
Db 5576 GCTTTCCTCTCTGACACATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 5635
QY 1315 IleSerGlnProGlyIleLeuHisProLeuValIleProSerValIleAlaSerThrPro 1334
Db 5636 ATTTCACAGCTGTGGGCTGCATCTTGTGTACTTCCATCATGATGATGCTTCTACTCTCT 5695
QY 1335 IleLeuProGlnAlaAlaGlyProThrSerThrProLeuLeuProGlnValProSerIle 1354

Db 5696 ATTTCCTCCAGAGAGAGGACTTACTTACACTTTATATCCCAAGTACCTAGATTC 5755
QY 1355 ProProLeuValGlnProValAlaAsnValProAlaValGlnIleThrLeuIleHisSer 1374
Db 5756 CCACCTTGGTAGACCTGTGTGCCAATGTGCTCTGTGACAGACAGACACTAATTCATAGT 5815
QY 1375 GlnProGlnProAlaLeuLeuProAsnGlnProHisThrHisCysProGlnValAspSer 1394
Db 5816 CAGCTCAACCGCTTGTGCTTCCACACAGCCCACTCACTCACTTCTCTGAAAGTGAATTC 5875
QY 1395 AspThrGlnProIleValProGlyIleAspAspIleIysThrLeuGlnGluIleLeuArg 1414
Db 5876 GATACACACCAACCAAGCTCTCGAATTTGATGACATTAAGATCTTAGAAGAAAGACTGGG 5935
QY 1415 SerLeuPheSerGluHisSerSerSerGlyAlaGlnHisAlaSerValSerLeuGluThr 1434
Db 5936 TCTCTGTTCAGTGAACAGCTCATCTGAGCTCAGATGCTCTGTCTCACTGGAGACC 5995
QY 1435 SerLeuValIleGlnSerThrValThrProGlyIleProThrThrAlaValAlaProSer 1454
Db 5996 TCACTAGTCAATAGAGACACTGTACACAGCAGCATCCCACTACTGCTGTGACACAGC 6055
QY 1455 LysLeuLeuThrSerThrThrThrSerThrCysLeuProProThrAsnLeuProLeuGlyThr 1474
Db 6056 AAACCTCTGACTTCAACACAGTACTTGTATACACCAACCATTTTACCTAGAAACA 6115
QY 1475 ValAlaLeuProValThrProValValThrProGlyGlnValSerThrProValSerThr 1494
Db 6116 GTTGTCTTTCAGTTCACACAGAGTGTCAACCTGGGCAAGTTTCTACCCAGTACAGCT 6175
QY 1495 ThrThrSerGlyValIleProGlyThrAlaProSerIysProProLeuThrIleValPro 1514
Db 6176 ACTACATCAGAGTGAAGAACTGTGAATCTCTCTCCAGCGCACTTCTTAATGAGCTTCG 6235
QY 1515 ValLeuProValGlyThrGluLeuProAlaGlyThrLeuProSerGluIleProPro 1534
Db 6236 GTGCTGCAGTGGGTACTGAATCTCCAGCAGATACCTTACCCAGCAGAGCTGCCACT 6295
QY 1535 PheProGlyProSerLeuThrGlnSerGlnProLeuGlnAspLeuAspAlaGlnLeu 1554
Db 6296 TTTCCAGAGACCTTCTCAACCCAGCTCCACAGCACTTCAAGATCTTGATGTCTCAATTTG 6355
QY 1555 ArgArgThrLeuSerProGluIleThrValThrSerAlaValGlyProValSerMet 1574
Db 6356 AGAAGAACACTTGTGTCAGATGATATACAGTATCTTCTCGGTGTGCTCTGTCATG 6415
QY 1575 AlaAlaProThrAlaIleThrGluAlaGlyThrGlnProGlnIleGlyValSerGlnVal 1594
Db 6416 GCGGCTCYAACACCAATCACRGAAGCAGAAACAAGCTCAGAAAGGTGTTCTCAAGTC 6475
QY 1595 LysGluGlyProValLeuAlaThrSerSerGlyAlaGlyValPheIleMetGlyArgPhe 1614
Db 6476 AAAGAAAGCCCTGTCTTACGAACTAGTTCAGAGCTGTGTGTTTTTAAGTGGAGCATTT 6535
QY 1615 GlnValSerValAlaAlaAspGlyAlaGlnIleGlyValIleGlyValIleGlyValIle 1634
Db 6536 CAGCTTCTGTGTCAGCAGACGCTGCCAGAAAGAGGCTTAATAATAGTACAGAGATGCA 6595
QY 1635 LysSerValHisPheGlnSerSerThrSerGluSerSerValLeuSerSerSerPro 1654
Db 6596 AAGCTGTCTATTGTGATTCACACACTCAGAGTCTCAGTGTCTATCAAGTATGATGCA 6655
QY 1655 GluSerThrLeuValIleProGluProArgGlyIleThrIleProGlyIleSerSerAsp 1674
Db 6656 GAGAGTACTTGTGTGAACAGAGCGCAATGGATACCACTTCTGTATCTTTCAGAT 6715
QY 1675 ValProGluSerAlaHisIleThrThrAlaSerGlyAlaIleSerSerAspThrGlyGlnPro 1694
Db 6716 GTGCCAGAGGTGCCCAAACTACTGCTCAGAGGCAAGTCAAGACTGTGGGAGCT 6775
QY 1695 ThrLysValGlyArgPheGlnValThrThrAlaAsnLysValGlyArgPheSerVal 1714

QY 121 ArgGluGluThrValThrAlaThrSerGluValAlaGlnGlnProProAlaIa 140
 Db 1316 CGGAGAGAGACCGGACCGCCACCTTCCAGAGAGCCAGACCTCCAGCCGCT 1375
 QY 141 AlaAlaProGlyGluGlnAlaValAlaGlyProAlaProSerThrValProSerSerThr 160
 Db 1376 RCGCCCGCCGGGAGAACAGGCGCGCGGGCCCTGCCCTCGACTGTCGCCAGCAGTACC 1435
 QY 161 SerIysAspArgProValSerGlnProSerLeuValGlySerIysGluGluProProPro 180
 Db 1436 AGCAAAAGCCCGCCAGAGTCCAGCTTAGCTTGCGGAGCAAGAGACCGCCCGCG 1495
 QY 181 AlaArgSerGlySerGlyGlyGlySerAlaIysGluProGlnGluGluValSerGlnGln 200
 Db 1496 GCGAGAGTGGCAGCGCGCGCGCGCCCAAGAGCCACAGAGAGAGAGAGCCAGCAG 1555
 QY 201 GlnAspAspIleGluGluLeuGluThrIysAlaValGlyMetSerAsnAspGlyArgPhe 220
 Db 1556 CAGAGTATATCGAAGAGCTGAGACCAAGCGCGGGAATGCTTAACGATGCCGCTTT 1615
 QY 221 LeuIysPheAspIleGluIleGlyArgGlySerPheIysThrValIleTyrGlyLeuAsp 240
 Db 1616 CTCAAGTTTGACATCGAAATCGGAGAGGCTCTTTAAGACGCTCAAAAGGCTCGAC 1675
 QY 241 ThrGluThrThrValGluValAlaIleTyrCysGluLeuGlnAspArgIysLeuThrIysSer 260
 Db 1676 ACTGAACCAACCGAGAAAGTGGCTGGTGTGAATCGCAGATCCAAATTTAACAAAGCTT 1735
 QY 261 GlnArgGlnArgPheIysGluGluAlaGluMetLeuIysGlyLeuGlnIleProAsnIle 280
 Db 1736 GAGAGGACGAGATTTAAAGAAAGAGCTGAATGTTAAAGCTTCACGATCCCAATATT 1795
 QY 281 ValArgPheTyrAspSerTyrGluSerThrValIysGlyIysCysIleValLeuVal 300
 Db 1796 GTTAGATTTTAGATCTCGGGAATCCACAGTAAAGAAAGAGCTGATGTTTGCTG 1855
 QY 301 ThrGluLeuMetThrSerGlyThrLeuIysThrTyrLeuIysArgPheIysValMetIys 320
 Db 1856 ACTGAACCTTAGACGCTCGAAGACCTTAAACGATCTGAAAGAGTTTAAAGTATGAG 1915
 QY 321 IleIysValLeuAspSerTyrCysArgGlnIleLeuIysGlyLeuGlnPheLeuIleThr 340
 Db 1916 ATCAAAAGCTTCAAAAGCTGCTGCGTCAATCTTAAAGGCTTCGATTCCTTCATACT 1975
 QY 341 ArgThrProProIleIleHisArgAspLeuIysCysAspAsnIlePheIleThrGlyPro 360
 Db 1976 CGAACTCACCTTATCATTCACCGGATCTTAAATGTACAAACATCTTATACCGGCTCT 2035
 QY 361 ThrGlySerValIysIleGlyAspLeuGlyLeuAlaThrLeuIysArgAlaSerPheAla 380
 Db 2036 ACTGGCTCAGTCAAGATTTGGAGACTCGGCTGCGCAACCCGGAAGCGGGCTCTTTGCC 2095
 QY 381 LysSerValIleGlyThrProGluPheMetAlaProGluMetIleTyrGluIysTyrAsp 400
 Db 2096 AAGGTGTGTAGTACCCAGAGTTCATGGCCCTCAGATGTATGAGGAGAAATATGAT 2155
 QY 401 GluSerValAspValTyrAlaPheGlyMetCysMetLeuGluMetAlaThrSerGluTyr 420
 Db 2156 GAATCGTTTACGTTTAGCTTTGGATGTGATGCTTGAGATGCTACCTCGAATAT 2215
 QY 421 ProTyrSerGluCysGlnAsnAlaIleGlnIleTyrArgArgValThrSerGlyValIys 440
 Db 2216 CTTACTCGAGTCCAAAATGCTGCGACAGATCTACCGTGGCGTGAACAGTGGGTAG 2275
 QY 441 ProAlaSerPheAspIysValAlaIleProGluValIysGluIleIleGluGlyCysIle 460
 Db 2276 CCGAGCCAGTTTTCGAAAGTAGCAATTCGAGATGAGGAAATATATGAAGATGATCA 2335
 QY 461 ArgGlnAsnIysAspGluArgTyrSerIleIysAspLeuLeuAsnIleAlaPhePheGln 480
 Db 2336 CGACAAAACAAAGATGAAAGTATTCATCAAAAGACTTTTGAACCAATGCTTTTCCAA 2395
 QY 481 GluGluThrGlyValArgValGluLeuAlaGluGluAspAspGlyGluIysIleAlaIle 500

Db 2396 GAGAAACAGAGTACCGGTGTAATTTAGCAGAGAAAGATGAGAAAATATACCTA 2455
 QY 501 LysLeuTyrLeuArgIleGluAspIleIysIysLeuIysGlyIysTyrIysAspAsnGlu 520
 Db 2456 AAATATGAGTACGTAATTTAAAGATATTTAAGAAATTTAAAGGAAAAATACAAAGATATCA 2515
 QY 521 AlaIleGluPheSerPheAspLeuGluArgAspValProGluAspValAlaGlnIleMet 540
 Db 2516 GCTATTAGATTTCTTTTGATTTAGAGAGATATCCAGAGATGTTGTCGCAAGAAATG 2575
 QY 541 ValGluSerGlyTyrValCysGluGlyAspHisIysThrMetAlaIysAlaIleIysAsp 560
 Db 2576 GTAGAGTCTGGGTATGTCTGTGAGGTGATACAAAGCATGCTTAAGCTATCAAAAGAC 2635
 QY 561 ArgValSerLeuIleIysArgIysArgGluGlnArgIleLeuValArgGluGluGln 580
 Db 2636 AGAGTATCATTTAATTAAGAGAAACGAGAGCAGCGGAGTTGTTACGGAGAGCAAGAA 2695
 QY 581 LysIleIysGlnGluGluSerSerLeuIysGlnGlnValGluGlnSerSerAlaSerGln 600
 Db 2696 AAAAAAAGCAGAAAGAGAGAGCTCAAAACAGAGGTAGAAACAATCCAGTCTCCAG 2755
 QY 601 ThrGlyIleIysGlnLeuProSerAlaSerThrGlyIleProThrAlaSerThrThrSer 620
 Db 2756 ACAGAGATCAAGCAGCTCCCTTCTGCTAGCAGCGGCATCTACTGCTTACCACTTCA 2815
 QY 621 AlaSerValSerThrGlnValGluProGluGluProGluAlaAspGlnIleGlnIleu 640
 Db 2816 GCTTCAGTTTCTACACAAGTAGAACTGAGAAACCTGAGGACGATCAACATCAACACTA 2875
 QY 641 GlnTyrGlnGlnProSerIleSerValLeuSerAspGlyThrValAspSerGlyGlnGly 660
 Db 2876 CAGTACAGCAACCACTATATCTGTGTATCTGATGAGGAGTGAACAGGTGACAGGA 2935
 QY 661 SerSerValPheThrGluSerArgValSerSerGlnGlnThrValSerTyrGlySerGln 680
 Db 2936 TCCTCTCTCTTCAVAGATCTCAGTGAAGCAGCAACAGCAAGTTTATATGTTCCCA 2995
 QY 681 HisGluGlnAlaHisSerThrGlyThrValProGlyHisIleProSerThrValGlnAla 700
 Db 2996 CATGAACAGCACATTCATTCAGGACAGCTCCAGGGCATATVACCTTACTGTCCAGCA 3055
 QY 701 GlnSerGlnProHisGlyValTyrProProSerSerValAlaGlnGlyGlnSerGlnGly 720
 Db 3056 CAGTCTAGCCCAAGGGGTATATCCACCTCAAGTGTGACAGGGGACAGCCAGAGCT 3115
 QY 721 GlnProSerSerSerSerLeuThrGlyValSerSerSerGlnProIleGlnIleProGln 740
 Db 3116 CAGCAATCCCTCAAGTACTTAAACAGGGGTTTCATCTTCCCAACCAACAAATCCTCAG 3175
 QY 741 GlnGlnGlnGlyIleGlnGlnThrAlaProProGlnGlnThrValGlnTyrSerLeuSer 760
 Db 3176 CAGCAGAGGAGAAATACAGCAGACAGCCCCCTTCMAACAGACAGTGAATTTCACTTCA 3235
 QY 761 GlnThrSerThrSerSerGluAlaThrThrAlaGlnProIleSerGlnProGlnAlaPro 780
 Db 3236 CAGACATCAACCTCCAGTGAAGGCCATCTACGACAGCAGCAGGATCARCCTCAAGCTCA 3295
 QY 781 GlnValLeuProGlnValSerAlaGlyIysGln 791
 Db 3296 CAAGTCTTGCTCAAGATACAGTGGAAACAGCTTCCAGTTTCCAGCAGTACCAACT 3355
 QY 791 ----- 791
 Db 3356 ATCCAAAGCGAAGCTCAGATCCAGTTGCGACAAACCTCGGTTGTTCCAGTCACTCT 3415
 QY 791 ----- 791
 Db 3416 GGTGCTCATTTCTTCCAGTGGAGACAGCGCTCCCTTATCTCCCTTGCTCCTCAGTACCT 3475
 QY 791 ----- 791

QY 1335 IleLeuProGlnAlaIleGlnProThrSerThrProLeuLeuProGlnValProSerIle 1354
Db 5696 ATTCTTCCCAAGCAGAGAGACTTCTTACACCTTTATTACCCAAAGTACTAGATC 5755
QY 1355 ProProLeuValGlnProValAlaAsnValProAlaValGlnInThrLeuIleHisSer 1374
Db 5756 CCACCTTGATGAGCTGTGGCCATGTGCTGCTGACAGAGACACTAATTCATAGT 5815
QY 1375 GlnProGlnProAlaLeuLeuProAsnGlnProHisThrHisCysProGlnValAspSer 1394
Db 5816 CAGCTTCACACAGCTTGTGCTCCACACAGCCCATCTCATTTGCTCGAAGTAGATTCT 5875
QY 1395 AspThrGlnProValAspProGlyIleAspAspIleThrThrLeuGlnGlyIleValArg 1414
Db 5876 GATACACACCCAAAGCTCTGGAATTGATGACATTAAGACTCTAGAGAAAGAACTCCG 5935
QY 1415 SerLeuPheSerGlnHisSerSerSerGlyAlaGlnHisAlaSerValSerLeuGlnThr 1434
Db 5936 TCTCTGTTCAAGTGAACACACCTCATCTGAGCTCAGCATCCTCTGCTCAGCTGAGACC 5995
QY 1435 SerLeuValIleGlnSerThrValThrProGlyIleProThrThrAlaValAlaProSer 1454
Db 5996 TCCTACTGATGATGAGACACCTGTCACACAGGACATCCCACTACTGCTTGACCAAGC 6055
QY 1455 LysLeuLeuThrSerThrThrSerThrCysLeuProProThrAsnLeuProLeuGlyThr 1474
Db 6056 AAACCTCTGACTTACCAACAAGTACTGCTTACCAACAACAATTATTCACATAGAAC 6115
QY 1475 ValAlaLeuProValThrProValValThrProGlyGlnValSerThrProValSerThr 1494
Db 6116 GTTGCTTTGACAGTTACACAGAGTGCACACCTGGGCAAGTTCTTACCCACAGTACAGCT 6175
QY 1495 ThrThrSerGlyValLysProGlyThrAlaProSerLysProProLeuThrLysAlaPro 1514
Db 6176 ACTACATCAGAGAGGAACTGGAACCTGCTCCCTCCAAAGCACCTTAACTAAAGCTCCG 6235
QY 1515 ValLeuProValGlyThrGlnLeuProAlaGlyThrLeuProSerGlnGlnLeuProPro 1534
Db 6236 GTGTGTCACTGGGTACTGAACTTCCAGAGTACTTACCCAGCGAGCTGCCACT 6295
QY 1535 PheProGlyProSerLeuThrGlnSerGlnGlnProLeuGlnAspLeuAspAlaGlnLeu 1554
Db 6296 TTTCACAGACTTCTCTAACCCAGTCCACAGCACTCTAGAGGATCTTGATGCTCAATTG 6355
QY 1555 ArgArgThrLeuSerProGln**IleThrValThrSerAlaValGlyProValSerMet 1574
Db 6356 AGAAGAACACTGTAGTCAGAGATKATACAGTAGCTTCTGGGTGTGCTGTGTCATG 6415
QY 1575 AlaAlaProThrAlaIleThrGlnAlaGlyThrGlnProGlnLysGlyValSerGlnVal 1594
Db 6416 GCGGCTCYAACAGCAATCACAGAACGAGAACACAGCTCAGAAAGGTGTTTCCAGTCTC 6475
QY 1595 LysGlnGlyProValLeuAlaThrSerSerGlyAlaGlyValPheLysMetGlyArgPhe 1614
Db 6476 AAAGAAAGCCCTGCTCAGCAACTAGTTCAGAGACTGTGTTTAAAGATGGGACGATTT 6535
QY 1615 GlnValSerValAlaAlaAspGlyAlaGlnLysGlnGlyLysAsnLysSerGlnAspAla 1634
Db 6536 CAGGTTCCTGTGAGAGAGAGGTGCCAGAAAGAGGTAAATAAGTCAAGAAAGATGCA 6595
QY 1635 LysSerValHisPheGlnSerSerThrSerGlnSerValLeuSerSerSerSerPro 1654
Db 6596 AAGTCTGTTCTATTGATCCAGCACTCAGAGTCTCAAGTGCATACAGTAGTAGTCCCA 6655
QY 1655 GlnSerThrLeuValLysProGlnProAsnGlyIleThrIleProGlyIleSerSerAsp 1674
Db 6656 GAGAGTACTTGTGGAACAGAGCCCAATGSCATTAACCATCCCTGGATACTCTTCAGAT 6715
QY 1675 ValProGlnSerAlaHisLysThrThrAlaSerGlnAlaLysSerAspThrGlnPro 1694
Db 6716 GTGCAAGAGAGTGGCCCAAAAATACTGCTCAGAGGCAAAAGTCACTCGGCAAGCTT 6775
QY 1695 ThrLysValGlyArgPheGlnValThrThrThrAlaAsnLysValGlyArgPheSerVal 1714
Db 6776 ACCAAGTTGAGCTTTTCAGCTGACACATACAGCAACAAAGGSGTCTTCTCTGTA 6835
QY 1715 SerLysThrGlnLysLysIleThrAspThrLysLysGlyGlyProValAlaSerProPro 1734
Db 6836 TCAAAAATCTGAGCAAGAATCACTGACACAAAGAAAGAAAGAACAGTGGCATCTCTCT 6895
QY 1735 PheMetAspLeuGlnGlnAlaValLeuProAlaValIleProLysLysGlyLysProGln 1754
Db 6896 TTTATGATTTGAAACAGCTGTTCTTCTGCTGTGATACAAAGAAAGAAAGCTTGA 6955
QY 1755 LeuSerGlnProSerHisLeuAsnGlyProSerSerAspProGlnAlaIlePheLeuSer 1774
Db 6956 CTCTCAGAGCTTCAATCAATTAATGGGCGCTCTTCTGACCCGAGGCGCTTTTAAAT 7015
QY 1775 ArgAspValAspAspGlySerGlySerProHisSerProHisGlnLeuSerSerLysSer 1794
Db 7016 AGGATGTGATGATGATGTTCCGGTAGTCCACACTGCCCATCAGCTGAGCTCAAAAGC 7075
QY 1795 LeuProSerGlnAsnLeuSerGlnSerLeuSerAsnSerPheAsnSerSerTyMetSer 1814
Db 7076 CTTCCTAGCCAGAAATTAAGTCAAAGCTTAGTATTATTAATTAATCTCTTACATAGT 7135
QY 1815 SerAspAsnGlnSerAspIleGlnAspGlnAspLeuLysLeuGlnLeuArgArgLeuArg 1834
Db 7136 AGCAGCAATGATGATGATATCGAAAGATGAAGACTTAAGTAAAGTGAAGTGCAGCACTCA 7195
QY 1835 AspLysHisLeuLysGlnIleGlnAspLeuGlnSerArgGlnLysHisGlnIleGlnSer 1854
Db 7196 GATTAACATCTCAAAGAGATTCAAGACCTGCAAGTGCAGAGACATGAAGAAATGAATCT 7255
QY 1855 LeuThrThrLysLeuGlnGlyLysValProProAlaValIleIleProProAlaIleProLeu 1874
Db 7256 TTGTATATCCAAACTGGGCAAGTGGCCCCCTGCTGTATTATTTCCCCAGCTGCTCCCTT 7315
QY 1875 SerGlyArgArgArgProThrLysSerLysGlySerLysSerSerArgSerSerSer 1894
Db 7316 TCAGGAGGAAGAGACAGACCCCACTAAAGCAAAAGCAGCAATCTAGTCCAAAGCAGTTCC 7375
QY 1895 LeuGlyAsnLysSerProGlnLeuSerGlyAsnLeuSerGlyGlnSerAlaIleSerVal 1914
Db 7376 TTGGGGAATTAAGACCCCGCTTCAAGTAACTGTGTGATCAGAGTGAAGCTTCAGTCTC 7435
QY 1915 LeuHisProGlnGlnThrLeuHisProProGlyAsnIleProGlnSerGlyGlnAsnGln 1934
Db 7436 TTGCACCCCAGAGACACCTTCCACCTCTGCGAACATCCAGAGTCCGGGAGAAATCAG 7495
QY 1935 LeuLeuGlnProLeuLysProSerProSerSerAspAsnLeuTySerAlaPheThrSer 1954
Db 7496 CTGTTACAGCCCTTAAGCCATCTCCCTCAGTAGACAACCTCTATTCAGCTTCCACAGT 7555
QY 1955 AspGlyAlaIleSerValProSerLeuSerAlaProGlyGlnGlyThrSerThrAsn 1974
Db 1956 GATGTGTCCATTTTCAGTACAGACCTTTCTGCTCCAGGTCAAGGAACACAGACACAAC 7615
QY 1975 ThrValGlyAlaThrValAsnSerGlnAlaIleGlnAlaGlnProProAlaMetThrSer 1994
Db 7616 ACTGTTGGGCAACAGTGAAGACAGCAAGCCGCCAACCTCAGCTCTGCAATAGACTCC 7675
QY 1995 SerArgLysGlyThrPheThrAspAspLeuHisLysLeuValAspAspThrAlaArgAsp 2014
Db 7676 AGCAGGAAGGAGCATTCACAGATGACTTGCACAAAGTGTAGACAAATTTGGGCCCGAGAT 7735
QY 2015 AlaMetAsnLeuSerGlyArgArgGlySerLysGlyHisMetAsnTyrgLysProGly 2034
Db 7736 GCCATGATCTTCAGAGAGAGAGAGAGAGCAAGAGGACATGAATTAAGAGGCGCTCGA 7795
QY 2035 MetAlaArgLysPheSerAlaProGlyGlnLeuCysIleSerMetThrSerAsnLeuGly 2054
Db 7796 ATGGCAAGGAAGTTCTTGCACCTGGGCACTGTGCATCTCCATGACTCGAAGCTGGGT 7855
QY 2055 GlySerAlaProIleSerAlaIleSerAlaThrSerLeuGlyHisPheThrLysSerMet 2074

QY 481 GlnGluThrGlyValArgValGluLeuAlaGluGluAspAspGlyGluValLeuAlaIle 500
 DB 2386 GAGGAAACGAGAGTACGGGTAAATTAAGCAAGAGATGAGAAATTAAGCCATA 2455
 QY 501 LysLeuThrPheArgIleGluAspIleLysLysLeuLysGlyLysThrLysAspAsnGlu 520
 DB 2456 AATATATGGCTACGATATTAAGATATTAAGAAATTAAGGAAATTAAGATATATGA 2515
 QY 521 AlaIleGluPheSerPheAspLeuGluArgAspValProGluAspValAlaGlnGluMet 540
 DB 2516 GCTATTGAGCTTTCTTTGATTAGAGAGATGTCCAGAAAGTGTCCACAGAAAG 2575
 QY 541 ValGluSerGlyLysValAlaCysGluGlyAspHisLysThrMetAlaLysAlaIleLysAsp 560
 DB 2576 GTAGAGCTGGGATGTCTGTGAAGGTGATCACAAACCAATGCTTAAGCTATCAAGAGC 2635
 QY 561 ArgValSerLeuIleLysArgLysArgGluGlnArgGlnLeuValArgGluGlnGlu 580
 DB 2636 AGAGTATCATTTAATTAAGGAAACGAGAGCGGAGGAGTGGTACGGGAGGAGCAAGAA 2695
 QY 581 LysLysLysGlnGluGluSerSerLeuLysGlnGlnValGluGlnSerSerAlaSerGln 600
 DB 2696 AAAAAAACAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2755
 QY 601 ThrGlyLysGlnLeuProSerAlaSerThrGlyLysProThrAlaSerThrThrSer 620
 DB 2756 AGGGAATCAAGAGATCTCTCTGTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2815
 QY 621 AlaSerValSerThrGlnValGluProGluGluProGluAlaAspGlnHisGlnGlnLeu 640
 DB 2816 GCTTCAAGTTTCAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2875
 QY 641 GlnThrGlnGlnProSerIleSerValLeuSerAspGlyThrValAspSerGlyGlnGly 660
 DB 2876 CAGTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2935
 QY 661 SerSerValPheThrGlnGlnSerArgValSerSerGlnGlnThrValSerThrGlySerGln 680
 DB 2936 TCTCTGTCTTCAAGATCTCAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2995
 QY 681 HisGlnGlnAlaHisSerThrGlyThrValProGluHisIleProSerThrValGlnAla 700
 DB 2996 CAGGAAACGAGAGATCTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3055
 QY 701 GlnSerGlnProHisGlyValLysProProSerSerValAlaGlnGlyGlnSerGlnGly 720
 DB 3056 CAGTCTCAGCCCATGGGTATATCCACCTCAAGTGTGACAGAGAGAGAGAGAGAGAGAG 3115
 QY 721 GlnProSerSerSerSerLeuThrGlyValSerSerSerGlnProIleGlnHisProGln 740
 DB 3116 CAGCCATCTCAAGTACCTTAACAGGGGTTTCACTTCCCAACCCATACAACTCTCTCAG 3175
 QY 741 GlnGlnGlnGlnLysGlnGlnThrAlaProProGlnGlnThrValGlnLysSerLeuSer 760
 DB 3176 CAGAGAGAGAGAGATACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3235
 QY 761 GlnThrSerThrSerSerGluAlaThrThrAlaGlnProValSerGlnProGlnAlaPro 780
 DB 3236 CAGACATCAACCTCAAGTGAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3295
 QY 781 GlnValLeuProGlnValSerAlaGlyLysGln----- 791
 DB 3296 CAAAGTCTGCTCAAGATACAGCTGAGAAACAGCTTCCAGTTTCCAGCCAGATACCAACT 3355
 QY 791 ----- 791
 DB 3356 ATCCAAAGCGAAGCTCAGATCCAGTTGCGACACAACCTCGGTTGTTCCAGTCCACTCT 3415
 QY 791 ----- 791
 DB 3416 GGTGCTCATTTCTTCCAGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3475
 QY 791 ----- 791

DB 3476 GGTCTCARTTCCCATATCAACTCTCATATGTGTACAGGCTCAGACAGGTTTCTCATCC 3535
 QY 791 ----- 791
 DB 3536 CTTCCCATCAAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3595
 QY 791 ----- 791
 DB 3596 ACAACAGCTGCAGATCCGGGGGTATCACTGTGTCTTGTAGCTTCCAAACCTTCTG 3655
 QY 791 ----- 791
 DB 3656 CAGCCTGTACTCAGGTGCAAGTCAAGTTCAACCAAGCTCTTCAACAGAGAGTTCCAG 3715
 QY 791 ----- 791
 DB 3716 TCCATGGGAATACAGCTAACCCTTGACAGAGCTGTGAGTTCCACTTCTCTGTGAGAT 3775
 QY 791 ----- 791
 DB 3776 GTTCTGTACAGAGGCTTCCCACTGACCTGACCAACAGATACCCAGAGATTCATATTT 3835
 QY 791 ----- 791
 DB 3836 GCTCCCTTCCAGAGTGGCTTCTGTGTCATCATTCTACAGTCCATCCCTCCCATG 3895
 QY 791 ----- 791
 DB 3896 CCGAGAAAGTACTGGCTACACCTGGGTACTTCCCAAGTGTGACAGCTTATGTGAA 3955
 QY 791 ----- 791
 DB 3956 TCATACTTTAGTTCTTATGGTGTGTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4015
 QY 792 -----SerThrGln 794
 DB 4016 GGGAGTTAGCAAGCCCCCACTACATCTCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4075
 QY 795 GlyValSerGlnValAlaProAlaGluProValAlaValAlaGlnProGlnAlaThrGln 814
 DB 4076 GAGGTCTCAGAGTGTCTCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4135
 QY 815 ProThrThrLeuAlaSerSerValAspSerAlaHisSerAspValAlaSerGlyMetSer 834
 DB 4136 CCGACCACTTGGCTTCTCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4195
 QY 835 AspGlyAsnGluAsnValProSerSerSerGlyArgHisGluGlyLysThrThrLysArg 854
 DB 4196 GATGCAATGAGAACTCCCATCTTCCAGTGAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4255
 QY 855 HisThrArgLysSerValArgSerArgSerArgHisGluLysThrSerArgProLysLeu 874
 DB 4256 CATTACCGAAATCTGTAAGAGAGTCTCTCGACATGAAATTAATTTACAGCCCAAAATTA 4315
 QY 875 ArgIleLeuAsnValSerAsnLysGlyAspArgValAlaGlnCysGlnLeuGlnThrHis 894
 DB 4316 AGAATTTTGAATCTTCAATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4375
 QY 895 AsnArgLysMetValThrPheLysPheAspLeuAspGlyAspAsnProGluGluIleAla 914
 DB 4376 AATAGGAAATGTTCATTCATCAATTTGACCTGTAGTGTACAAACCCGAGGAGATGCA 4435
 QY 915 ThrIleMetValAsnAsnAspPheIleLeuAlaIleGluArgLysSerPheValAspGln 934
 DB 4436 ACAATTTATGTGAACAATGACTTATTTCTGCAATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4495
 QY 935 ValArgGluIleIleGluLysAlaAspGluMetLeuSerGluAspValSerValGluPro 954
 DB 4496 GTGCGAAGAAATTTATGAAAAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4555
 QY 955 GlnGlyAspGlnGlyLeuGlnSerLeuGlnGlyLysAspAspLysGlyLysPheSerGlySer 974

D	b		4556	SAGGGTGTATCAGGATTGGAGAAGCTCAACAGAAAGATGCATATGCTTTACAGTTCT	4615
O	y		975	GlnIysLeuGIuGIyLpHelySGLnProIIeProAlaSerSetMetProGInGlnIle	994
D	b		4616	CAGAAATATGGAAGAGAGATTCCAAACAACCATTCCTGGCTCTTCATGCCACAGAAATA	4675
O	y		995	GlyIleProThrSerSerLeuThrGlnValHisSerAlaGlyValArgPheIleVal	1014
D	b		4676	GGCATTTCTCACCAATCTTTTACTCAACTGATTTCTTCTCGGGAGAGCGGTTAATAGTG	4735
O	y		1015	SerProValProGluSerArgLeuArgLysSerIysValPheProSerGlnIleThrAsp	1034
D	b		4736	AGCTCTGGCCAGAAAGCCGATTCAGCAAAATCCAAGATTTTCCCAGGAATAACGAT	4795
O	y		1035	ThrValAlaAlaSerThrAlaGlnSerProGlyMetLeuSerHisSerAlaSerSer	1054
D	b		4796	ACAGTTGGTGGCTCTCAACGCTCAGAGCCCTGGAAATGAATCTGTCTCACTTCGATCATCC	4855
O	y		1055	LeuSerLeuGInGlnAlaPheSerGluLeuArgAlaGlnMetThrGluGIyProAsn	1074
D	b		4856	CTTAGTCTAACAAACAGGCTTTTTCGAATTAGACGTGCCCAAATGACAGAGACCAGAN	4915
O	y		1075	ThrAlaAProProAsnPheSerHisethrGlyProthrPheProValValProProPheLeu	1094
D	b		4916	ACAGACCTCCAAAATTGATCATACAGACCAACATTTCCAGTAGAACCTCCCTTCTTA	4975
O	y		1095	SerSerIleAlaGlyValProThrThrAlaAlaThrAlaProValProAlaThrSer	1114
D	b		4976	AGTAGCATTCGTCGAGTCCCAACACACACACACACACACACACAGTCCCTGCACCAAGC	5035
O	y		1115	SerProProAsnAspIleSerThrSerValIleGlnSerGluValThrValProThrglu	1134
D	b		5036	AGCCCTCCTAATGACATTTCCACATCAGTAATTCAGTCTGAGTTACAGTSCCACTGAA	5095
O	y		1135	GluGlyIleAlaGlyValAlaAlaThrSerThrGlyValValThrSerGlyGlyLeuProIle	1154
D	b		5096	GAGGGGATGCTGGAGTTGCCACACACACACAGAGTGTGGTACCTTCAGGGGTCTCCCATTA	5155
O	y		1155	ProProValSerGluSerProValLeuSerSerValValSerSerIleThrIleProAla	1174
D	b		5156	CCACTGTGTCTGAATCCACGACTTTTCCAGCGTAGTTTAAAGTATCACATACCTGCA	5215
O	y		1175	ValValSerIleSerThrThrSerProSerLeuGlnValProThrSerThrSerGluIle	1194
D	b		5216	GTGTGTCAATATCTACTACATCCCCGCTCACTTCAAGTCCCCACATCACAATCGAATC	5275
O	y		1195	ValValSerSerThrAlaLeuTyProSerValThrValSerAlaThrSerAlaSerAla	1214
D	b		5276	GTTGTTCTTAGTAAAGCACTGTATCTTCACATTAACAGTTTCAGCAACTTCAGCTCTGCA	5335
O	y		1215	GlyGlySerThrAlaThrProGlyProIlysProProAlaValValSerGInGlnAlaAla	1234
D	b		5336	GGGGGAGTACTGCTACCCCAAGTCTTAGCGCTTCAAGCTGTAGTATCTACAGAGGACGA	5395
O	y		1235	GlySerThrThrValGlyAlaAlaThrLeuThrSerValSerThrThrThrSerPheProSer	1254
D	b		5396	GGCAGCACTACTGTCGGAGCAACATTAACATCAGTTTCTACACACACTCATTTCCAGAC	5455
O	y		1255	ThrAlaSerGlnLeuSerIleGlnLeuSerSerThrSerThrProthrLeuAlaGln	1274
D	b		5456	ACACCTTCACAGCTGTSCATTCACGTTTAGCAGCAGTACTTCTACTCTTACTTTCAGTAA	5515
O	y		1275	ThrValValValSerAlaHisSerLeuAspIysThrSerHisSerSerThrThrGlyLeu	1294
D	b		5516	ACCGTGTAGTTTAGGACACTCACTCAATATAGACATTCACACACAGTACCAACTGATGG	5575
O	y		1295	AlaPheSerLeuSerAlaProSerSerSerSerSerProGlyAlaGlyValSerSerTyr	1314
D	b		5576	GCTTTTCCCTCTCGACCAATCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	5635
O	y		1315	IleSerGlnProGlyGlyLeuHisProLeuValIleProSerValIleAlaSerThrPro	1334
D	b		5636	ATTTCCTACGCTGGGGCTGCATCTTGGTCAATTCATCAGATAGATAGCTTCACTACCTT	5695

QY	1335	lleuPProGlnAlaIaAGlyProThrSerThrProLeuLeuProGlnValProSerIle	1335
Db	5656	ATTCTTCCCAAGACAGACGAGCACTTCTTACACCTTTATTACCCCAAGTACCTGATAC	5755
QY	1355	ProProLeuValGlnProValAlaAsnValProAlaValGlnGlnThrLeuIleHisSer	1374
Db	5756	CCACCTTGGTATACAGCCTGTTGCCAATGTGCTGCTGTATACAGACAGACACTAATTCATATGT	5815
QY	1375	GlnProGlnProAlaLeuLeuProAsnGlnProHisThrHisCysProGlnValAspSer	1394
Db	5816	CAGCCTCACACAGCCTTGCTTCCCAACAGCCCACTACTCATTTGCTCGAAGTGTGATCTT	5875
QY	1395	AspThrGlnProCysAlaProGlyTlIeAspAspIleCysThrLeuGlnGlnValLeuArg	1414
Db	5876	GATACACACACCAAGCTCTCTGGAATTTGATACATTAAAGACTCTTAGAAGAAAGTGGCG	5935
QY	1415	SerLeuPheSerGlnHisSerSerSerGlyValaGlnHisAlaSerValSerLeuGlnThr	1434
Db	5936	TCTCTGTTCAAGTAAACACAGCTCATCTGTGAGCTCAGCATGCGCTGTCTCACTGGAAGAC	5955
QY	1435	SerLeuValIleGlnSerThrValThrProGlyIleProThrThrAlaValAlaProSer	1454
Db	5996	TCACAGTCATAGAGAGACCTGTACACACAGGCAATCCCACTACTGCTGTGACACCAAGC	6055
QY	1455	LysLeuLeuThrSerThrThrSerThrCysLeuProProThrAsnLeuProLeuGlyThr	1474
Db	6056	AAACTCGTCACTTCTACACCAAGTACTGCTTACACCAACCAATTTTACCACTAGAACAA	6115
QY	1475	ValAlaLeuProValThrProValValThrProGlyGlnValSerThrProValSerThr	1494
Db	6116	GTTGGTTTCCAGATTAACCAAGTGTGTACACCTGAGGAAGTTTTCATCCCGAGTACGACT	6175
QY	1495	ThrThrSerGlyValIlyProGlyYThrAlaProSerIysProProLeuThrIlySalPro	1514
Db	6176	ACTCATATAGSAGTGAACCTCTGGAATCTCTCCCTCCAAAGCACCCTTACTTAAGGCTCG	6235
QY	1515	ValLeuProValGlyYThrGlnLeuProAlaGlyYThrLeuProSerGlnGlnLeuProPro	1534
Db	6236	GTGCTGCAGAGGGTACTGAACCTTCCAGAGGTACTCTACCCACGACGAGCTCCACCT	6295
QY	1535	PheProGlyProSerSerLeuThrGlnSerGlnGlnProLeuGlnAspLeuAspAlaGlnLeu	1554
Db	6296	TTCACGACCTTCTCTACCCCAATCCAGCAACCTCTAGAGGATCTTGATGCTCTCAATG	6355
QY	1555	ArgArgThrLeuSerProGlu**IleThrValThrSerAlaValGlyProValSerMet	1574
Db	6356	AGAAGAACAACCTTAGTCCAGAGATATATACAGTACTTCTGGGTTGGTCTGTCTCATG	6415
QY	1575	AlaAlaProThrThrAlaIleThrGlnAlaGlyYThrGlnProGlnIlyGlyValSerGlnVal	1594
Db	6416	GCGGCTCYAAAGCAATCACGGAAGCAAGAAACAAGCCTCAAGAGGGTGTTCCTCAAGTC	6475
QY	1595	LysGlnGlyProValIleLeuAlaThrSerSerSerGlyAlaGlyValPheIysMetGlyArgPhe	1614
Db	6476	AAAGAAGCCCTGTCTCTAGCAACTAGTTCAGGAGCTGGTGTTTTAAGATGGACGATTT	6535
QY	1615	GlnValSerValAlaAlaAspGlyValaGlnIlyGlnIlyLysAsnIysSerGlnAspAla	1634
Db	6536	CAGGTTCTCTGTTCAGACAGACGGTGGCCAGAAAGCGTAAATAATAGTACAGAAGATGCA	6595
QY	1635	LysSerValHisPheGlnSerSerThrSerGlnSerSerValLeuSerSerSerSerPro	1654
Db	6536	AAAGTCTGTTCAATTGTAATCCAGCACTTCAGAGTCTCTCAGTGCTATTAAGTAGTAGTCCA	6655
QY	1655	GlnSerThrLeuValIlyProGlnProAsnGlyIleThrIleProGlyIleSerSerAsp	1674
Db	6656	GAGAGTACCTTGGTGAACCAAGACCGCAATGAGTAAACATCCCTGTGATCTCTTCAGAT	6715
QY	1675	ValProGlnSerAlaHisIlyThrThrAlaSerGlnAlaLysSerAspThrGlnPro	1694
Db	6716	GTTGCAGAGATGCCCCAATAACTACTGCTTCAGAGGCAAGATGACACTCTGGGCAAGCTT	6775

Db 1256 TCCTGCCCCAGCCAGCATCCCGCGCTGTCCCGAGAGTGTCTCCAGCCGAGACCCAC 1315
QY ArgGluGluThrValThrAlaThrAlaThrSerGlnValAlaGlnGlnProProAla 140
Db 1316 CGGAGAGAGACCGGAGACCGCCACCTTCCAGGTAGCCGACGAGCCCTCCAGCCGCT 1375
QY AlaAlaProGlyGlnGlnAlaValAlaGlyProAlaProSerThrValProSerSerThr 160
Db 1376 RCCCCCTCGGGGAAACAGCGCGTGGCGGCGCTCCCTCCGACGTGTCCAGCAGTACC 1435
QY SerLysAspArgProValSerGlnProSerLeuValGlySerLysGlnGlnProProPro 180
Db 1436 AGCAAGAGCGCCAGAGTGTCCAGCTTGTGTGGAGCAAAAGAGAGCCGCCCGC 1495
QY AlaArgSerGlySerGlyGlySerAlaLysGlnProGlnGlnGlnArgSerGlnGln 200
Db 1496 GCGAGAGTGGACGCGCGCGGAGCGCCAGAGAGCCACAGAGAGAGAGAGCCACAG 1555
QY GlnAspAspIleGlnGlnLeuGlnThrLysAlaValGlyMetSerAspAspGlyArgPhe 220
Db 1556 CAGGATGATTCGAAAGAGTGGAGACCAAGGCGTGGGAATGTCTAACGATGGCGCTTT 1615
QY LeuLysPheAspIleGlnIleGlyArgGlySerPheLysThrValTyrLysGlyLeuAsp 240
Db 1616 CTCAAGTTTACATCGAAATCGGACAGAGCTCTTTAAGACGGTCTCAAAAGGCTGGAC 1675
QY ThrGluThrThrValGluValAlaTyrProCysGlnLeuGlnAspArgGlySerLeuThrLysSer 260
Db 1676 ACTGAAACCAACCGGAGAGTGGCTGGTGTGACGTGACGAGATCGAAATTAACAAGTCT 1735
QY GlnArgGlnArgPheLysGlnGlnAlaGlnMetLeuLysGlyLeuGlnHisProAsnIle 280
Db 1736 GAGAGCGAGAGATTAAAGAAAGAGCTGAATGTTAAAGGTCTTCAGCATCCCAATATT 1795
QY ValArgPheLysAspSerTyrGlnSerThrValLysGlyLysGlyIleValLeuVal 300
Db 1796 GTTAGATTATATGATTCCTGGGAATCCACAGTAAAGAAAGAGTCAATGTGTTGGTG 1855
QY ThrGluLeuMetThrSerGlyThrLeuLysThrTyrLeuLysArgPheLysValMetLys 320
Db 1856 ACTGAACTTATGAGCTGTGAGAACCTTAAACGATCTGAAAGGTTTAAAGTATAG 1915
QY IleLysValLeuArgSerTyrProCysArgGlnIleLeuLysGlyLeuGlnPheLeuHisThr 340
Db 1916 ATCAAACTTCTAAAGAGCTGGTGGCTGACATCTTAAAGTCTTCAGTTTCTTCACT 1975
QY ArgThrProProIleIleHisArgAspLeuLysCysAspAsnIlePheIleThrGlyPro 360
Db 1976 CGAACTCACCTTATCATTCACCGGATCTTAAATGTACAAACATCTTATCACCGGCCCT 2035
QY ThrGlySerValLysIleGlyAspLeuGlyLeuAlaThrLeuLysArgAlaSerPheAla 380
Db 2036 ACTGGCTCACTCAAGATGGAGACCTGGTCTGGCAACCTGAAAGCGGCTTCTTTGGC 2095
QY LysSerValIleGlyThrProGlnPheMetAlaProGlnMetTyrGlnGlnLysTyrAsp 400
Db 2096 AAGAGTGTGATGATGCCAGAGTTCATGAGCCCTCAGATGTATGAGAGAAATATGAT 2155
QY GlnSerValAspValTyrAlaPheGlyMetCysMetLeuGlnMetAlaThrSerGlyTyr 420
Db 2156 GAATTCGTACGTTATGCTTTTGGGATGTGATGCTTAAGATGGCTAATCTGAATAT 2215
QY ProTyrSerGlnCysGlnAsnAlaAlaGlnIleTyrArgArgValThrSerGlyValLys 440
Db 2216 CTTTACTCGGAGTGCAGAAATGCTGCRCAATCTACCGTGGCTGACCACTGGGGTATAG 2275
QY ProAlaSerPheAspLysValAlaIleProGlnValLysGlnIleIleGlnGlyCysIle 460
Db 2276 CCGAGCCAGTTTTCACAAAGTAGCAATTCGAAAGTAGAGAAATTAATTGAAGATGATCA 2335
QY ArgGlnAsnLysAspGluArgTyrSerIleLysAspLeuLeuAsnHisAlaPhePheGln 480
Db 2336 CGACAAACAAAGATGAAAGATATTCATCAAGAGCTTTTGAACCATGCTTCTTTCAA 2395

QY 481 GlnGluThrGlyValArgValGluLeuAlaGlnLysAspAspGlyGlnLysIleAlaIle 500
Db 2396 GAGAAACAGAGATGACGGTATGAAATTCAGAAAGAGAGATGAGAGAAAAAATATGCCATA 2455
QY LysLeuTyrLeuArgIleGlnAspIleLysLysLeuLysGlyLysTyrLysAspAsnGln 520
Db 2456 AAATTATGCTTACGTATTAAGATATTAAGAAATTAAGGAAATTAACAAGATATCA 2515
QY AlaIleGlnPheSerPheAspLeuGluArgAspValProGlnMetPheValAlaGlnGlnMet 540
Db 2516 GCTATTGAGTTTCTTTTGTATTAAGAGAGATGTCCCAAGATGTGTGACAAAGAAATG 2575
QY ValGlnSerGlyTyrValCysGlnGlyAspHisLysThrMetAlaLysAlaIleLysAsp 560
Db 2576 GTNAGCTTGGATATGTCTGTGAAGTATCAACAGCCATGGCTTAAGCTATCAAGAGC 2635
QY ArgValSerLeuLysArgLysArgGlnGlnArgGlnLeuValArgGlnGlnGlnGln 580
Db 2636 AGAGTATCATTAATTAAGAGAAACAGAGACAGCGGAGTTGGTACGGAGAGCAAGAA 2695
QY LysLysLysGlnGlnGlnGlnSerSerLeuLysGlnGlnValGlnGlnSerAlaSerGln 600
Db 2696 AAAAAAAGCAGAGAGAGACAGTCTCAACAGAGGTAGAACATCCAGTGTCTCCAG 2755
QY ThrGlyLysGlnLeuProSerAlaSerThrGlyLysProThrAlaSerThrThrSer 620
Db 2756 ACAGAGATCAAGAGCTCTCTTGTGTCAGACCGGCTACTACTGCTTCAACACTTCA 2815
QY AlaSerValSerThrGlnValGlnProGlnGlnProGlnLysAspGlnHisGlnGlnLeu 640
Db 2816 GCTTCAGTTTCTACACAGTAGAACCAGAAACCTGAGGACAGATCAACATCAACACTA 2875
QY GlnTyrGlnGlnProSerIleSerValLeuSerAspGlyThrValAspSerGlyGlnGly 660
Db 2876 CAGTACAGCAACCCAGTATATCTGTATCTGATGGAGCGGTGACAGTGTGCAGAGA 2935
QY SerSerValPheThrGlnSerArgValSerSerGlnGlnThrValSerTyrGlySerGln 680
Db 2936 TCCTGTCTTTCAYAGATTCAGATGACAGACCAAGCAGACAGTTCATATGTGCTCCA 2995
QY HisGlnGlnAlaHisSerThrGlyThrValProGlyHisIleProSerThrValGlnAla 700
Db 2996 CATGAACAGGACATTTCTACAGGACAGTCCAGGGCATATACCTTACTATGTCCAAGCA 3055
QY GlnSerGlnProHisGlyValTyrProProSerSerValAlaGlnGlnGlnSerGlnGly 720
Db 3056 CAGTCTCAGCCCATGGGTATATCCACCTCAAGTGTGGCAAGGGGAGAGCCAGAGGT 3115
QY GlnProSerSerSerSerLeuThrGlyValSerSerSerGlnProIleGlnHisProGln 740
Db 3116 CAGCATCTCCTCAGATGACTTAACAGGGGTTTCACTTCCCAACCATCAACATCTCCAG 3175
QY GlnGlnGlnGlnIleGlnGlnThrAlaProProGlnGlnThrValGlnTyrSerLeuSer 760
Db 3176 CAGCAGACAGGAGATPACAGCAGACGCCCTCTCAACAGACAGTACGATATTCACCTTCA 3235
QY GlnThrSerThrSerSerGlnAlaThrThrAlaGlnProValSerGlnProGlnAlaPro 780
Db 3236 CAGACATCAACCTCCAGTGGAGGCGCCTACTACGACAGCAGATGAGTCAACCTCAAGCTCA 3295
QY GlnValLeuProGlnAlaSerAlaGlyLysGln----- 791
Db 3296 CAGTCTTGCTCAAGTATAGCTGAGAAACAGCTTCAAGTTCCAGCCAGTACCAACT 3355
QY ----- 791
Db 3356 ATCCAAAGCGAAGCTCAGATCCAGTTGCGACACAAACCTCGGTTGTTCCAGTCACTCT 3415
QY ----- 791
Db 3416 GGTGCTCATTTCTTCCAGTGGAGACGCGGCTCCCTATCTCCCTTGTCCCTCAGTACCT 3475

QY	791	-----	791
Db	3476	GTCCTCARTATCCCATATCAATCCTCGATGTGTTACAGGCTCAGACAGATTTCATCC	3535
OY	791	----	791
Db	3536	CTTCCCATCACATGGCAGCTGGGACTTACTGACGCTGTGCTCAGGTGGCTTCACTGCT	3593
OY	791	-----	791
Db	3596	ACAACAGCTGGGATCCGGGGGTATCAACTGTGTTCTTAGTGACATTCCAACCCTCTG	3655
OY	791	-----	791
Db	3656	CAGCCTGTGACTCAGCTGCCAAGTCAGTTCCACCAGCTCTTAACAACCAGAGTTGAG	3715
OY	791	-----	791
Db	3716	TTCATGGGAATAACGACTAACCTTBSGACAAGCTGTGAGTTCCACTTCTCTGGAGAT	3776
OY	791	-----	791
Db	3776	GTTCTGTACACAGGCTTCCCACCTGCAGTGCACACAGTACCAGAGATTCAAATATT	3835
OY	791	-----	791
Db	3836	GCTCCCTCTTCCACGTGGCTTCTGTTTGCATTCATCTACAGTCTATCCCCCTCCATG	3895
OY	791	-----	791
Db	3896	CCGACAGAAGTACGTGGCTACACACTGGGTACTTCCACAGTGGTGCAGCCTTATGTGAA	3955
OY	791	-----	791
Db	3956	TCAAATCTTTAGTCTCTATGGGTGGTGAAGAGACAGTTCAAGTGTCCAGCCAGGA	4015
OY	792	-----SerThrgIn	794
Db	4016	GGGAGTTTAGCACAGCCCCCACTACATCTCTCCAGCAAGCAGTTTGGAGTGACTCAG	4075
OY	795	GIyAlSerGIInValAlaProIagIuProValAlaValAlaGlInProGIInAlarhGIN	814
Db	4076	GGAGTCTTCAGGTTGCTCTGTCGAGGCCAGTTGAGTACAGACAGMCCCAAGCTACCGAG	4133
OY	815	ProThrThrlenuAlaSerSerValAspSerAlahISerAspValAlaSergIyMetSer	834
Db	4136	CCGACCACTTTGGCTTCTCTGTGTAGACAGTGACATTCAGATGTTGCTTCAGATGAGT	4195
OY	835	AspGIyAnGIuASrValProSerSerSergIyAlghISgluGIyArGrThrltySarG	854
Db	4196	GATGGCAATYGAAAGTCCCATCTTCCAGTGGAGGCAATGAAAGGAACAATAAAAACGG	4255
OY	855	HISrYrArgIySSerValArgSerArgSerArgHISgluIySThSerArgProIySLen	874
Db	4256	CATTRCCGAAATCTGTAAAGAGTGCCTCGACACTGAAAAACTTCCAGCCCCAAATTA	4315
OY	875	ArgIIeIeuASnValSerASnLySGIyASpArgValValGIuCySgInleugIuthrHis	894
Db	4316	AGAAATTTTGAATGTTTCAATATPAAGAGACCGAGTAGTAGATGCATTTAGAGACTCAT	4375
OY	895	ASNArgIyMeVAlThPhelyPheAspleuAspGIyASPsnProGIuSlulealA	914
Db	4376	AATGGGAAAAAGGTACATTCAMATTTTACCTAGATGTGTACMAACCCGAGAGATACGA	4435
OY	915	ThrIleMetValASnASnAPheIIeIeuAlaleGIuArGrGIuSerPheValAspGIIn	934
Db	4436	ACAAATTTATGTGAACAATGACTTTATTTCTAGCAATATGAGAGAGACTCGTTGTGATCAA	4495
OY	935	ValARgIuIlleIeGIuLyAlaAspGIuMetLeuSerGIuASpValSerValGIuPro	954
Db	4496	GTCGGAGAAATATTATGAAAAAGCTGATTAAGTCCATGTAGAGAGTCAAGTGTGAAACA	4555
OY	955	GIuGIyAspGIInGIyLeuGIuSerIeuGIuInGIyLysAspAspTyrgIyPheSergIySer	974

Db	4556	GAGGGTATACAGGGATTGGAGAGCTTACAGGAAAGGATACATAGGCTTTTCAGGTTCT	4615
Oy	975	GLInylSerLugInylGluPheLysGlnProIleProAlaSerSerMetCProGlnGlnIle	994
Db	4616	CAGAAATTTGAGAGAGAGGTTCAACAACACAAATTCCTGCTTCCTCCAGCCAGCAATATA	4675
Oy	995	GLylLeProThrSerSerLeuThrGlnValValHisSerAlaGlyValArgPheIleVal	1014
Db	4676	GGCATTCCTTACAGATTCCTTTAACTCAAGTGGTTCATTCCTCGGGAAGCGGTTATAGTG	4735
Oy	1015	SeerProValProGluSerArgLeuArgGlnSerLysValPheProSerGlnIleThrAsp	1034
Db	4736	AGTCTGTGGCAGAAAGCCGATTTACGGAATTCAAAAGTTTTCCTCCAGGAAATACAGAT	4795
Oy	1035	ThrValAlaAlaSerThrAlaGlnSerProGlyMetAsnLeuSerHisSerAlaSerSer	1054
Db	4796	ACACTTCTGCTCTTACAGCTCAGAGGCCCTGGATGAACTTGTTCTCATCTGCATCATCC	4855
Oy	1055	LeuSerLeuGlnGlnAlaPheSerGlnLeuArgArgAlaGlnMetThrGluGlyProAsn	1074
Db	4856	CTTAGTCTTACAAACAGGGCCTTTTTCGAATTAAGCTGCCAAATGACAGAAAGGCCAAN	4915
Oy	1075	ThrAlaProProAsnPheSerHisThrGlyProThrPheProValValProProPheLeu	1094
Db	4916	ACACACCTCCAAACTTTAGTCATACAGAGCAACAATTTCCAGTAGTAACTCTTCTTTA	4975
Oy	1095	SeerSerIleAlaGlyValProThrThrAlaAlaAlaThrAlaProValProAlaThrSer	1114
Db	4976	AGTAGCATTCGTGGAGTCCCACTACACAGACGACCCACAGCACCAAGTCCCTGCACAAAGC	5035
Oy	1115	SeerProProAsnAspIleSerThrSerValIleGlnSerGlnValThrValProThrGln	1134
Db	5036	AGCCCTCTTATATGACATTTCCACATCAGTAATTCAGTCTAGGTTTACAGTGCCACATGAA	5095
Oy	1135	GlnGlyIleIleAlaGlyValAlaThrSerThrGlyValValThrSerGlyLeuProIle	1154
Db	5096	GAGGGAGATGCTGGAGTGGCCACACGACGAGGTGTGAATCTTCAGGTGTCTCCCATATA	5155
Oy	1155	ProProAlaSerGlnSerProValLeuSerSerValValSerSerIleThrIleProAla	1174
Db	5156	CCACTGTGTCTGATACACCAAGTACTTTCCAGCTGATTTTCAAGTATCACAAATACCTGCA	5215
Oy	1175	ValValSerIleSerThrThrSerProSerLeuGlnValProThrSerThrSerGlnIle	1194
Db	5216	GTGTCTCATATATCTACTACATCCCTCCCTCATCTCAATCCCAATCATCATCTGAAATGC	5275
Oy	1195	ValValSerSerThrAlaLeuTyrProSerValThrValSerAlaThrSerAlaSerAla	1214
Db	5276	GTGTATTTCTAGTACAGCACTGTATCTCTTCAAGTAAACATTTCCAGCACTTCAGACTTCGCA	5335
Oy	1215	GlyValSerThrAlaThrProGlyProLysProProAlaValValSerGlnGlnAlaAla	1234
Db	5336	GGGGGAGTACTGCTTACCCACAGGTCCTTAAGCCTCCAGCTGATAGTATCTCAGGAGCGACA	5395
Oy	1235	GlySerThrThrValGlyAlaThrLeuThrSerValSerThrThrThrSerPheProSer	1254
Db	5396	GGCAGCACTACTGCGGAGCGACATTAACTACATGATTTCTTACCAACCACTTCATTCACAGC	5455
Oy	1255	ThrAlaSerGlnLeuSerIleGlnLeuSerSerThrSerThrProThrIleAlaGln	1274
Db	5456	ACACCTTCCACAGCTGTCSCATTCACCTTACAGACAGTACTTACTTCTACTTAAAGCTGAA	5515
Oy	1275	ThrValValAlaSerAlaHisSerLeuAspLysThrSerHisSerSerThrThrGlyLeu	1294
Db	5516	ACCGTGTAGTACGCGACACTCACTGATTAAGACATCTCATACAGATCAACTGCATTTG	5575
Oy	1295	AlaPheSerLeuSerAlaProSerSerSerSerSerProGlyAlaGlyValSerSerTyr	1314
Db	5576	GCTTTTCCCTCTCTGACACATCTTCTTCTCTCTGAGAGAGAGGTCTAAGTTAT	5635
Oy	1315	IleSerGlnProGlyGlyLeuHisProLeuValIleProSerValIleAlaSerThrPro	1334

Db	2336	CGACAAAGAAAGATGAAAGATATTCATCAAAAGACCTTTGAAACCATGCTTTTCCAA	2395
QY	481	GIUGIuThrGlyValArgValGluLeuValGluGluAspGlyGlyIleValIle	500
Db	2396	GAGGAAACAGAGATACGGGTAGAAATTAGCAGAAAGAGATGAGGAAAAATAGCCATA	2455
QY	501	IysLeuTPLeuuhrgTLeGluuhpTLeysIysLeuysGlyLysTyLysAspAsnIu	520
Db	2455	AAATTATGGCTACGATATGAAATGATTTAAAGAAATTTAAAGGGAATAATCAAGATATGAA	2515
QY	521	AlaIleGluPheSerPheAspLeuGluArgAspValProGluAspValAlaGluIleuMet	540
Db	2516	GCTATTGAGCTTTCTTTTGATTTAAGAGAGAGTGTGCCAAGAGAGTGTGCAAGAAATG	2575
QY	541	ValGluSerGlyTyTyrValCysGluGlyAspHisIysThrMetAlaLysAlaIleLysAsp	560
Db	2576	GTAGAGTCTGGGTATGCTGTGAAGTATTCACAGACCAATGCTTAAGGCTTATCAAAAC	2635
QY	561	ArgValSerLeuTLeysArgLysArgGluGluArgGluLeuValArgGluGluGlnIu	580
Db	2636	AGAGTATCATTTAATTAAAGGAAACGAGAGCAGCGGACGTTGTGATCGGAGAGAGAA	2695
QY	581	LysIysLysGluGluGluSerSerLeuysGlnGluValGluGluSerSerAlaSerGln	600
Db	2696	AAAAAAAAGCAGAAAGAGAGAGCTCTAAACAGCAGGTAGAACATTCAGAGCTTCCAG	2755
QY	601	ThrGlyIleLysGluLeuProSerAlaSerThrGlyIleProThrAlaSerThrThSer	620
Db	2756	ACAGAAATCAAGCAGCTCCCTTCTGCTAGCAGCGGCATACCTACTGCTTCAACACTTCA	2815
QY	621	AlaSerValSerThrGluValGluProGluGluProGluAlaAspGlnHisGluGlnLeu	640
Db	2816	GCTTCAGATTTCTACACAGTAGAAACCTGAAACCTGAGCAGATCAACATCAACAACATA	2875
QY	641	GlnTyGlnGlnProSerTLeSerValLeuSerAspGlyThrValAspSerGlyGlnGly	660
Db	2876	CAGTTCACAGCAACCCAGTATATCTGTGTTACTATGAGGAGGTTTGCACAGTGTCAAGGCA	2935
QY	661	SerSerValPheThrGluSerArgValSerSerGlnGlnThrValSerTyGlySerGln	680
Db	2936	TCCCTGTGCTTCATAGAAATCTCGAGTAGCAGCCAAACAGACGTTTCAATATGTTCCAA	2995
QY	681	HisGluGlnAlaHisSerThrGlyThrValProGlyHisIleProSerThrValGlnAla	700
Db	2996	CATGAACAGGCACTTCTACAGGACAGCAGTCCCAAGGCATATACTTCACTGTCAGAACCA	3055
QY	701	GlnSerGlnProHisGlyValTyTyrProProSerSerValAlaGlnGlyGlnSerGlnGly	720
Db	3056	CAGTCTCAGCCCACTGAGGTATATCCACCCCTCAAGTGTG-----	3094
QY	721	GlnProSerSerSerSerLeuThrGlyValSerSerSerGlnProIleGlnHisProGln	740
Db	3094	-----	3094
QY	741	GlnGlnGlnGlyIleGlnGlnThrAlaProProGlnGlnThrValGlnTySerLeuSer	760
Db	3095	-----CAGGGAATAACAGACAGACAGCCCTCTCAACAGACAGTACGATTTCACTTCA	3148
QY	761	GlnThrSerThrSerSerGlyAlaThrThrAlaGlnProValSerGlnProGlnAlaPro	780
Db	3149	CAGACATCAACTCTCAGTAGAGGCCACATCTGCACACGACAGTAGTCAACCTCAAGCTCCA	3208
QY	781	GlnAlaLeuProGlnValSerAlaGlyLysGlnSerThrGlnGlyValSerGlnValAla	800
Db	3209	CAATCTTCTGCTCAAGTATCACTGCGGAAAAACAGATATCTACAGGAGAGCTCTCAGGTTGCT	3268
QY	801	ProAlaGluProValAlaValAlaGlnProGlnAlaThrGlnProThrThrLeuAlaSer	820
Db	3269	CTTCAGACAGCAGATTGCAGTAGCACAACACCAAGCTTCCACAGCGCAACACTTTGGCTTCC	3328
QY	821	SerValAspSerAlaHisSerAspValAlaSerGlyMetSerAspGlyAsnGluAsnVal	840
Db	3329	TCGTATACACAGTGCACATTCAGATGTTGCTTCAAGTATGATGATGCAATGAGCAATGTC	3388

QY	841	ProSerSerSerSgLVaRgH;eGluLVaRgThrThrLysArgH;eStyArgLysSerVal	860
Db	3389	CCATCTTCCATGGGAAGGCAATGAAAGAACTCAAAAAGGCAATTACCGAAAACCTGTA	3448
QY	861	ArgSerArgSerArgH;eGluLVeThrSerArgProLysLeuArgLLeuAsnValSer	880
Db	3449	AGGATCGCTCGACATGAAAAAACTTCAGGCCCAAAATTAAAGATTTTGATCTTTCA	3508
QY	881	AsnLysGlyAspArgValValGluCysGluLeuGluThrH;eAsnArgLysMetValThr	900
Db	3509	AATAAAGAGACCGAGTGTAGAAATGTCATAATTAGACTCATTAATAGAAAAATGGTTACA	3568
QY	901	PhelysPheAspLeuAspGlyAspAsnProGluGluLeuAlaThrLLeuValAsnAsn	920
Db	3569	TTCAAAATTGACCTGATGCTGACAAACCCCAAGAGATGACAAATTATGCTGAACAT	3628
QY	921	AspPheLLeuAlaLLeuLysArgGluSerPheValAspGluValArgGluLeuLeu	940
Db	3629	GACCTTAATTCTAGCAATGAGAGAGAGAGTCGTTTGATGATCAAGTCGGAATAATTGAA	3688
QY	941	LysLAspGlyMetLeuSerGluAspValSerValGluProGluGlyAspGluGlyLeu	960
Db	3689	AAAGCTGATGAATGCTCAGTGAAGATGCACTGTGAACCAAGAGGGGATCAGAGATTG	3748
QY	961	GluSerLeuGluGlyLysAspAspTrpGlyPheSerGlySerGluLysLeuGluGlu	980
Db	3749	GAGAGCTTACAAAGAAAGAGATGACTATGGCTTTCAAGTTCTCGAAATTGGAAAGAG	3808
QY	981	PhelysGluProLLeuProLLeuSerMetProGluGluLeuGlyLeuProThrSerSer	1000
Db	3809	TTCAAAACCAACAATTCCTGCGCTCTTCATGCAACAGCAAAATAGCCATTCTCCACAGTTCT	3868
QY	1001	LeuThrGluValValH;eSerLArgArgPheLLeuValSerProValProGluSer	1020
Db	3869	TTAATCTCAAGTTGTTCAATCTGCGGGAAGGCGGTTTATAGTGAATCTGTGCGCAAGAGC	3928
QY	1021	ArgLeuArgGluSerLysValPheProSerGluLeuThrAspThrValAlaAlaSerThr	1040
Db	3929	CGATTACAGAAATCAAAAAGTTTCCCAAGTAATACAGATACAGTTGCGCTTACA	3988
QY	1041	AlaGluSerProGlyMetAsnLeuSerH;eSerLAserSerLeuSerLeuGluGluAla	1060
Db	3989	GCTCAGACCCCTGGAAATGAATCTTCTCAGCTCTGCAATCCTCTTAATCTTCAACAAGACC	4048
QY	1061	PheSerGluLeuAspArgAlaGluMetThrGluGlyProAsnThrAlaProProAsnPhe	1080
Db	4049	TTTTCTGACTTAAAGCTGCCAAATACAGAAGAGACCAAMNCAGACCTCCAAACTTT	4108
QY	1081	SerH;eThrGlyProThrPheProValValProProPheLeuSerLLeuAlaGlyVal	1100
Db	4109	AGTATACAGGACCAACATTTCCAGTGTACCTCTCTTTAAGTAGACATTGCTGAGAGTC	4168
QY	1101	ProThrThrAlaAlaAlaThrAlaProValProLAsThrSerSerProProAsnAspLe	1120
Db	4169	CCAACVACAGACAGGCCACAGACCAAGCTCCCTGCAACAAAGACCCCTCTTAATGAACAT	4228
QY	1121	SerThrSerValLLeuGluSerGluValThrValProThrGluGluGlyLeuAlaGlyVal	1140
Db	4229	TCCCATCAGTAATTCAGTCTGAGGTACAGTGCCCACTGAAGGGGAGTTGCTGGAGTT	4288
QY	1141	AlaThrSerThrGlyValValThrSerGlyGlyLeuProLLeuProProValSerGluSer	1160
Db	4289	GCCACCAAGCAAGGTGTGTAACTTCAAGGTGTCTCCCACTACCACTGTGTGAATCA	4348
QY	1161	ProValLeuSerSerValValSerSerLLeuThrLLeuAlaValValSerLLeuSerThr	1180
Db	4349	CCAGTACTTTCCAGCGTAAGTTTCAAGATACAAATCTGCAATTCCTGCAATATCTACT	4408
QY	1181	ThrSerProSerLeuGluValProThrSerThrSerGluLeuValValSerSerThrAla	1200
Db	4409	ACATCCCCCGTCACTTCAAGTCCCCCAATCAATCTGAGATCTGTTCTTTCTAGTACAGA	4468

1201 LeuYrProSerValThrValSerAlaThrSerAlaSerAlaGlySerThrAlaThr 1220
4469 CTGATACCTTCAGTAACAGTTTCAGAACTTCAGCTCTGCAAGGGGCGATCTGCTTACC 4528
1221 ProGlyProIleSerProAlaValIleSerGlnAlaAlaGlySerThrValIle 1240
4529 CCAAGCTCTAAGCTCCAGCTGTATGATCTCAGACGAGCAGCAGCACTACTGTGGGA 4588
1241 AlaThrLeuThrSerValSerThrThrThrSerPheProSerThrAlaSerGlnLeuSer 1260
4589 GCCACATTTAAACATCAGTTTCTACACCACTTCATTCACAGCAACGCTTCACGCTGTGC 4648
1261 IleGlnLeuSerSerSerThrThrThrProThrLeuAlaGlnThrValValIleSerAla 1280
4649 ATTCAAGTTACGACGAGTACTTCTACTCTTACTTACTTACTTACTTACTTACTTACTT 4708
1281 HisSerLeuAspLysThrSerHisSerSerThrThrGlyLeuAlaPheSerLeuSerAla 1300
4709 CACTCACTAGATAAAGACATCTCATAGCAGTAAACAGTGGCTTCTCTCTCTGCA 4768
1301 ProSerSerSerSerSerProGlyAlaGlyValSerSerThrIleSerGlnProGlyIle 1320
4769 CCAATCTTCTCTCTCTCTCTGAGCAGGAGTGTCTATATTTCTCAGCCTGTGGG 4828
1321 LeuHisProLeuValIleProSerValIleAlaSerThrProIleLeuProGlnAlaAla 1340
4829 CTGCATCTCTTTGGTCACTTCATCAGTATGCTTCTTCTCTTCTTCTTCTTCTTCTTCT 4888
1341 GlyProThrSerThrProLeuLeuProGlnValProSerIleProProLeuValGlnPro 1360
4889 GGAACCTACTTCTTCACTTATTAACCCAGTACTGATCCAGTATCCACCTTGGTATGAGCCT 4948
1361 ValAlaAsnValProAlaValGlnGlnThrLeuIleHisSerGlnProGlnProAlaLeu 1380
4949 GTTGCATATGTGCTGCTGTACAGCAGACACTAATTCATAGTCAGCCTCAACAGCTTGG 5008
1381 LeuProAsnGlnProHisThrHisSerProGlyValAlaSerSerAspThrGlnProLysAla 1400
5009 CTTCACCAACGAGCCCACTACTCATTTGCTCTGAACTGATCTGATACCAACCCAAAGCT 5068
1401 ProGlyIleAspAspIleLysThrLeuGlnGlnLysLeuArgSerLeuPheSerGlnHis 1420
5069 CCGGATATGATGACATTAAGACTCTGAAAGAAAGAGTGGCTCTGTTCAAGTAACAC 5128
1421 SerSerSerGlyAlaGlnHisAlaSerValSerLeuGlnThrSerLeuValIleGlnSer 1440
5129 AGCTCATCTGAGTCAAGTCACTGCTCTGTCTCACTGAGACCTCACTAGCATAGAGAGC 5188
1441 ThrValThrProGlyIleProThrThrAlaValAlaProSerLysLeuLeuThrSerThr 1460
5189 ACTGTCAACACGAGCATCCCACTACTGCTGTGTGCAACAGCAAACTCTGACTTCTACC 5248
1461 ThrSerThrCysLeuProProThrAsnLeuProLeuGlyThrValAlaLeuProValThr 1480
5249 ACAAGTACTTGTCTTACACCAACCAATTTACACTAGAAACAGTGTGTTGGCAGTTACA 5308
1481 ProValValThrProGlyGlnValSerThrProValSerThrThrThrSerGlyValLys 1500
5309 CCAAGTGTCAACCTGGGCAAGTTTCAACCCAGTCACTACTCATATCAAGGAGTAA 5368
1501 ProGlyThrAlaProSerLysProProLeuThrLysAlaProValIleProValGlyThr 1520
5369 CTTGGAACTGCTCTCTTCAAGCCACTCTTAATTAAGGCTCGGAGTGTGCGCAGTGGTACT 5428
1521 GlnLeuProAlaGlyThrLeuProSerGlnGlnLeuProPheProGlyProSerLeu 1540
5429 GAATTCACAGAGGTAATCTTACCCAGCAGAGCAGTGCACCTTTTCCAGGACCTTCTCTA 5488
1541 ThrGlnSerGlnGlnProLeuGlnAspLeuAspAlaGlnLeuArgGlnThrLeuSerPro 1560
5489 ACCCAAGTCCCAAGCAACTTGAAGATCTTGATCTCAATTGAAACAAACCTTAGTCCA 5548
1561 Gln**IleThrValThrSerAlaValGlyProValSerMetAlaAlaProThrAlaIle 1580

5549 GAGATATCAACGTGACTTCTGCGGTGGTCCGTGCTCCATGAGCGGCTCAACAGCAATC 5608
1581 ThrGlnAlaGlyThrGlnProGlnLysGlyValSerGlnValIleGlnGlyProValIleu 1600
5609 ACRGAAGCAGGAACACAGCTCAGAAAGGAGTGTTCATCAAGTCAAAAGGCGCTGTCTTA 5668
1601 AlaThrSerSerGlyValAlaGlyValPheLysMetGlyArgPheGlnValSerValAlaAla 1620
5669 GCACTAGTTCAAGAGCTGTGTGTTTTTAAGATGGACGATTTTACGTTCTGTGGACGA 5728
1621 AspGlyAlaGlnLysGlnLysLysAsnLysSerGlnAspAlaLysSerValHisPheGln 1640
5729 GACGGTCCCAAGAAAGAGGTAAATAAGTCAGAAAGTCAAAAGCTGTTCAATTTTGA 5788
1641 SerSerThrSerGlnSerSerValLeuSerSerSerSerProGlnSerThrLeuValLys 1660
5789 TCCAGACCTCAGAGTCTCAGTCTATCAAGTATGATTCAGAGGTACTGTGTGAA 5848
1661 ProGlnProAsnGlyIleThrIleProGlyIleSerSerAspValProGlnSerAlaHis 1680
5849 CCAAGCCCAATGGCATTAACATCTCTGATCTTTCAGATGTCCAGAGAGTGCAC 5908
1681 LysThrThrAlaSerGlnAlaLysSerAspThrGlyGlnProThrLysValGlyArgPhe 1700
5909 AAAACTAGTCCCTCAGAGCAAAAGTCAGACCTGGCAGCCTTCAAGGTGGACGTTTT 5968
1701 GlnValThrThrThrAlaAsnLysValGlyArgPheSerValSerLysThrGlnLysAspLys 1720
5969 CAGGTACCACTACACCAACCAAAAGTGGGTCTGTTCTCTGTATCAAAACCTGAGCAG 6028
1721 IleThrAspThrLysLysGlnGlyProValAlaSerProPheMetAspLeuGlnGln 1740
6029 ATCACTGACCAAAAGAAAGAGCAGTGGCATCTCTCTCTTTTATGATTTGGAAACA 6088
1741 AlaValLeuProAlaValIleProLysGlyLysGlyProGlnLeuSerGlnProSerHis 1760
6089 GCTGTTCTTCTGCTGTGATACCAAAAGAAAGAACCTGACCTGACAGAGCTTTCACAT 6148
1761 LeuAsnGlyProSerSerAspProGlnAlaAlaPheLeuSerArgAspValAspAspGly 1780
6149 CTAAATGGCGCTTCTCTGACCCGAGGCGCTTTTAAAGTAGAGTGTGAGATGAGT 6208
1781 SerGlySerProHisSerProHisGlnLeuSerSerLysSerLeuProSerGlnAsnLeu 1800
6209 TCCGATGATCAACATCGCCCACTCACTGATGCTCAAAAGCCTTCTTACCGAATCTA 6268
1801 SerGlnSerLeuSerAsnSerPheAsnSerSerLysSerSerAspAsnGlnSerAsp 1820
6269 AGTCAAGCCTTAGTAATTCATTTAACTCTCTTACATGATGAGCGCAATGAGTCAAGT 6328
1821 IleGlnAspGlnAspLeuLysGlnLeuArgLeuArgLysAspLysHisIleLeuGln 1840
6329 ATCGAAGATGAACCTTAAAGTTAAGCTGCGAGCTGAGATGAATTAACATCTCAAGAG 6388
1841 IleGlnAspLeuGlnSerArgLysGlnLysGlnIleGlnLeuSerLeuThrLysLeuGly 1860
6389 ATTCAGAGACTGCAAGATGCCAGAGACATGAATGAATCTTGTATCAACAACTGGGC 6448
1861 LysValProProAlaValIleIleProProAlaAlaProLeuSerGlyArgArgArg 1880
6449 AAGGTGCCCCCTGTTATTAATTCCTCCAGCTGCTCTCTTCAAGGAGAAAGCGAGA 6508
1881 ProThrLysSerLysGlySerLysSerSerArgSerSerSerLeuGlnLysAsnLysSerPro 1900
6509 CCCACTAAAGCAAGGACCAATCTTAGTGAAGCAGTTCTTGGGAAATTAAGCCCC 6568
1901 GlnLeuSerGlyAsnLeuSerGlyGlnSerAlaAlaSerValLeuHisProGlnGlnThr 1920
6569 CAGCTTTCAGGTAACTGTCTGTGTGAGTGAAGTCAAGCTTCAAGCTTGGACCCCGAGCAGCC 6628
1921 LeuHisProProGlyAsnIleProGlnSerGlyGlnAsnGlnLeuLeuGlnProLeuLys 1940

```

Db      6629 CTCACCCCTCTGGGACATCCAGAGTCCGGGACGAATCAGCTGTACAGCCCTTAAG 6688
QY      1941 ProSerProSerSerAspAsnLeuTyrSerAlaPheThrSerAspGlyAlaIleSerVal 1960
Db      6689 CCATCTCCCTCCAGTGCACAACTCTATTACGCTTCACCAAGTATGGTGGCATTTACAGTA 6748
QY      1961 ProSerLeuSerAlaProGlyGlnGlyThrSerSerThrAsnThrValGlyAlaThrVal 1980
Db      6749 CCAAACCTTTCTGCTCCAGGCTCAAGAACACACACACACAAACACACTGTGGGGCAACAGTGG 6808
QY      1981 AsnSerGlnAlaAlaGlnAlaGlnProProAlaMetCtnrSerSerArgIleGlyLthrPhe 2000
Db      6809 AACACCCAAAGCCGCCCAAGCTCAGCTCTCTGCGCATGAGTCCAGAGGAGGAGGACATTC 6868
QY      2001 ThrAspAspLeuHisLysLeuValAspAsnTrpAlaArgAspAlaMetAsnLeuSerGly 2020
Db      6869 ACAGATGACTTGCAACAAGTTGGTAGCAATTGGGCCCGAGATGCCATGATCTCTCAAGC 6928
QY      2021 ArgArgGlySerLysGlyHisMetAsnTyrGlnGlyProGlyMetAlaArgLysPheSer 2040
Db      6929 AGGAAGAGAGAGGAAAGGGGACATGATTAAGAGGGCCCTGGAAATGGCAAGAAAGTTCTCT 6988
QY      2041 AlaProGlyGlnLeuCysIleSerMetThrSerAsnLeuGlyGlySerAlaProIleSer 2060
Db      6989 GCACCTGGGGCAACTGTGATCTCCATGACCTCGAACCCTGGAGGCTGTGGCCCCCATCTCT 7048
QY      2061 AlaAlaSerAlaThrSerLeuGlyHisPheThrLysSerMetCysProProGlnGlnTyr 2080
Db      7049 GCAGCATCAGCTACCTCTCTAGGTCACTTCAACCAAGCTTAATGGTCCCCCACAGCAGTAT 7108
QY      2081 GlyPheProAlaThrProPheGlyValaGlnTrpSerGlyThrGlyGlyProAlaProGln 2100
Db      7109 GGCCTTCAGCTACCCCAATTTGGCGCTCAATGAGACTGGAGCGGGGGCCGACACACAG 7168
QY      2101 ProLeuGlyGlnPheGlnProValGlyThrAlaSerLeuGlnAsnPheAsnIleSerAsn 2120
Db      7169 CCACCTGGCCAGTTCACACTGCTGGGAACCTGCCTCTTTCAGAAATTTCAACATCAGCAAT 7228
QY      2121 LeuGlnLysSerIleSerAsnProProGlySerAsnLeuArgThrThr 2136
Db      7229 TTGCAGAAATCCATCAGCAACCCCAAGCTCCAACTCGGGAGCCACT 7276

RESULT 9
US-10-767-471-452
; Sequence 452, Application US/10767471
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: RHEUMATOID ARTHRITIS, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01505
; CURRENT APPLICATION NUMBER: US/10/767,471
; CURRENT FILING DATE: 2004-01-30
; NUMBER OF SEQ ID NOS: 50231
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 452
; LENGTH: 7853
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(7853)
; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1-
US-10-767-471-452

```

Alignment Scores:

Pred. No.:	0	Length:	7853
Score:	10567.50	Matches:	2096
Percent Similarity:	98.13%	Conservative:	0
Best Local Similarity:	98.13%	Mismatches:	11
Query Match:	97.74%	Indels:	29
DB:	7	Gaps:	1

US-10-010-720-14 (1-2136) X US-10-767-471-452 (1-7853)

OY	1	MetSerGIyGIyAlaAlaGIuLysGIuSerSerThrProGIySerLeuPheIuSerPro	20
Db	956	ATGCTGGCGGGCGCGCGAGAGAGAGACACTCCCGTTCCTGCTCTCGGCC	1015
OY	21	ProAlaProAlaProLyAsnGIySerSerSerSerValGIyGIuLysLeuGIy	40
Db	1016	CCGGCTCTGGCCCCCAAGATGGCTCCAGCTCCGATTCCTCCGTGGGGAGAACTGGGA	1075
OY	41	AlaAlaAlaAlaAspAlaValThrGIyArgThrGIuGIyIyArgArgArgArgHisThr	60
Db	1076	GCCGGGCGCGGAGCTGTGACCCGGAGAGACCGAGGGGTCAAGCGCGCGCACACT	1135
OY	61	MetAspLyAspSerArgGIyAlaAlaAlaAlaThrThrThrThrThrThrGluHisArgPhePhe	80
Db	1136	ATGGACAAAGACAGCCCGTGGGGCGCGCGACCACTACCACTGAGACACCGCTTCTTC	1195
OY	81	ArgArgSerValIleCysAspSerAsnAlaThrAlaLeuGIuLeuProGIyLeuProIeu	100
Db	1196	CGCCGAGACGTCATCTGTGACCTCCAAATGCCACTGCAGTGAAGCTTCCGGCGCTTCTCT	1255
OY	101	SerIeuProGIuProSerIleProAlaAlaProGIuSerAlaProProGIuProHis	120
Db	1256	TCCCTGCCCAAGCCAGCATCCCGCGGCTGTCCCGAGATGTCTCCACGGAAKCCCAAC	1315
OY	121	ArgGIuGIuThrValThrAlaThrAlaThrSerGIuValaGIuGIuProProAlaAla	140
Db	1316	CGGGAGAGAGACGTGACCGGCACCGCCACTTCCCAAGTAGCCAGACAGCTCCAGCGCT	1375
OY	141	AlaAlaProGIyGIuGIuAlaValAlaGIyProAlaProSerThrValProSerSerThr	160
Db	1376	RCCGCGCCCTGGAGAAACAGCCGTGCGGGGCCCTCCCTCTGACTGTCCACAGATACC	1435
OY	161	SerLyAspArgProValSerGIuProSerIeuValGIySerLyGIuGIuProProPro	180
Db	1436	AGCAAAAGCCGCCAGTCTCCCACTAGCTTGTGGGGAGCAAAAGAGAGCCGCCGCG	1495
OY	181	AlaArgSerGIySerGIyGIySerAlaLysGIuProGIuGIuArgSerGIuGIu	200
Db	1496	GCGAGAGGTGCACGCGGCGGCGAGCGCCAAAGAGCAACGAGGAAACGAGCACAG	1555
OY	201	GIuAspAspIleGIuGIuLeuGIuThrValAlaAlaGIyMetSerAsnAspGIyArgPhe	220
Db	1556	CAGATGATATCGAAGAGCTGGAGAACAGGCCGCGGGAAATGCTAACGATGGCCGCTT	1615
OY	221	LeuLyPheAspIleGIuIleGIyArgGIySerPheLyThrValIyLyGIyLeuAsp	240
Db	1616	CTCAAGTTGACATCGAAATCGGCAGAGGCTCCTTAAAGCGGTCTCAAAAGTCTGAC	1675
OY	241	ThrGIuThrThrValGIuValAlaITyPysGIuLeuGIuAspArgLySLeuThrLysSer	260
Db	1676	ACTAAACCAACCGGAGAGTGCCTGTGTGAATCGACGATCCAAAATTAAACAAAGCT	1735
OY	261	GIuArgGIuArgPheLySGIuGIuAlaGIuMetLeuLySGIyLeuGIuHisProAsnIle	280
Db	1736	GAGAGCGAGACTTTAAAGAAAGAGCTGAATGTTTAAAGGTCTTCAGCATCCAAATTT	1795
OY	281	ValArgPheLyAspSerITyPGLySerThrValLysGIyLysLyCysIleValLeuVal	300
Db	1796	GTTAGATTTTATGATTCCTCGGAAATCCACAGTAAAGAAAGAGTGCATGTTTGGTG	1855
OY	301	ThrGIuLeuMetThrSerGIyThrLeuLySerThrTyrtLeuLyAspPheLyValMetLys	320
Db	1856	ACTAACTTATGACGTCTGAAACCTTAAACGATCTGAAAGAGTTTAAAGTATGAG	1915
OY	321	IleLyValLeuArgSerITyPysArgGIuIleLeuLySGIyLeuGIuPheLeuHisThr	340
Db	1916	ATCAAAAGTTCTAAGAGCTGGTGCCTCAAGTCTTAAAGGTCTTCAGTTTCTTCATACT	1975
OY	341	ArgThrProProIleIleHisArgAspLeuLyCysAspAsnIlePheIleThrGIyPro	360
Db	1976	CGAACTCCACTATCATTCACCGGATCTTAAATGTGAACAACATCTTATACCGGCGCT	2035

QY 361 ThrGlySerValIysIleGlyAspLeuGlyLeuAlaThrLeuLysArgAlaSerPheAla 380
 DB ACTGGCTCAAGTTCAGAGTTCCTGCTGCAACCTGAAGCGGGCTTTTGGCC 2095
 QY 381 LysSerValIleGlyThrProGluPheMetAlaProGluMetTyrGluGluTyrAsp 400
 DB AAGAGTGTGATAGTACCCAGAGTTTCATGCGCCCTGAGATGTATGAGAGAAATATGAT 2155
 QY 401 GluSerValAspValTyrAlaPheGlyMetCysMetLeuGluMetAlaThrSerGluTyr 420
 DB GAATCCGTTACGTTATGCTTTGGATGTGATGCTTGAAGTGTACATCTGAATAT 2215
 QY 421 ProTyrSerGluCysGlnAsnAlaIleGlnIleTyrArgArgValThrSerGlyValIys 440
 DB CCTTACTCGAGTCCAAAATGCTGCKACAGATCTACCGTCCGCTGACCAGTGGGGTAA 2275
 QY 441 ProAlaSerPheAspLysValAlaIleProGluValLysGluIleIleGluGlyCysIle 460
 DB CCAAGCCAGTTTTCACAAAGTACCAATTCCTGAAGTGAAGAAATTTATGAAGATGCATA 2335
 QY 461 ArgGluAsnLysAspGluArgTyrSerIleLysAspLeuMetLysAlaPhePheGln 480
 DB CGACAAAACAAAGATGAAGATATTCATCAAGACCTTTGAACCATGCTTTTCCAA 2395
 QY 481 GluGluThrGlyValArgValGluLeuAlaGluGluAspAspGlyGluTyrIleAlaIle 500
 DB GAGGAACACAGAGTACCGGTGAGATTCAGACAGAAAYGATGAGAAATAATGCCATA 2455
 QY 501 LysLeuTyrLeuArgIleGluAspIleLysLysLeuLysGlyTyrTyrLysAspAsnGlu 520
 DB AAATTATGCGCTACGATATTGAAGATATTAGAAATTTAAAGGGAAATTTCAAAAGATTAAGAA 2515
 QY 521 AlaIleGluPheSerPheAspLeuGluArgAspValProGluAspValAlaGlnGluMet 540
 DB GCTATTAGATTTCTTTGATTAGAGAGATGTCCAGAGAGTGTGCAAGAAATG 2575
 QY 541 ValGluSerGlyTyrValCysGluGlyAspHisLysThrMetAlaLysAlaIleLysAsp 560
 DB GTAGAGTCTGGGTATGCTCTGGAAGTGAATCAACAGACATGCTTAAGCTATCAAAAGAC 2635
 QY 561 ArgValSerLeuIleLysArgLysArgGluGlnArgGluLeuValArgGluGluGln 580
 DB AGAGTATCATTAATTAGAGAGAAACGAGACGCGAGTGTGTAACGGAGAGCAAGAA 2695
 QY 581 LysIleLysGlnGluGluSerSerLeuLysGlnGlnAlaGluGlnSerSerAlaSerGln 600
 DB AAAAAAAGCAGAGAGAGAGAGCTCAACACAGAGGTAGAACAAATCCAGTGTCCAG 2755
 QY 601 ThrGlyIleLysGlnLeuProSerAlaSerThrGlyIleProThrAlaSerThrThrSer 620
 DB ACGAGATCAAGCAGCTCCCTTCTGCTAGACACCGGCATACCTTCTGCTTACCACTTCA 2815
 QY 621 AlaSerValSerThrGlnValGluProGluGluProGluAlaAspGlnHisGlnGlnLeu 640
 DB GCTTCACATTTCTACACAGTAGAACCTGAAGAACCTAGCGAGATCAACATCAACAATA 2875
 QY 641 GlnTyrGlnGlnProSerIleSerValLeuSerAspGlyThrValAspSerGlyGlnGly 660
 DB CAGTACACGACCAACGATATCTGTGTATCTATGAGGACGAGTGTGACAGTGGTCAAGGGA 2935
 QY 661 SerSerValPheThrGluSerArgValSerSerGlnGlnThrValSerTyrGlySerGln 680
 DB TCCTCTGCTTCAAGAAATCTCGAGTAGACGCCAACAGRCAGTTTCATATGTTCCCAA 2995
 QY 681 HisGlnGlnAlaHisSerThrGlyThrValProGlyHisIleProSerThrValGlnAla 700
 DB CATGAACAGGACATTTCTACAGGACAGTCCACAGGACATATACCTTCTACTGCTCAAGCA 3055
 QY 701 GlnSerGlnProHisGlyValTyrProProSerSerValAlaGlnGlnGlnSerGlnGly 720
 DB CAGTCTAGCCCATGGGGTATATCCACCTCAAGTGG----- 3094
 QY 721 GlnProSerSerSerSerLeuThrGlyValSerSerSerGlnProIleGlnHisProGln 740

DB 3094 ----- 3094
 QY 741 GlnGlnGlnGlyIleGlnGlnThrAlaProProGlnGlnGlnThrValGlnTyrSerLeuSer 760
 DB 3095 -----CAGGAATATACGACGACAGACCCCTCCATCAACAGACAGTGTATTCATCTTCA 3148
 QY 761 GlnThrSerThrSerSerGluAlaThrThrAlaGlnProValSerGlnProGlnAlaPro 780
 DB CAGACATCAACCTCCAGTGAAGGCCACTACTGTCACAGCCAGTGTACTCCTCAAGCTTCA 3208
 QY 781 GlnValLeuProGlnValSerAlaGlyLysGlnSerThrGlnGlyValSerGlnValAla 800
 DB CAAGTCTGCTCAAGATACAGTCAAGTGAAGAAACAGAGTACTCAGGAGAGTCTCTGAGTTGCT 3268
 QY 801 ProAlaGluProValAlaValAlaGlnProGlnAlaThrGlnProThrThrLeuAlaSer 820
 DB CCTGCAGAGCCAGTGTGACATGACAGCCAGCCAGCTAACCCAGCCGACCTTGGCTTCC 3328
 QY 821 SerValAspSerAlaHisSerSerValAlaSerGlyMetSerAspGlyAsnGluAsnVal 840
 DB TCTGTACACGTGACATTCAGATGTTGCTTCAGATGATGATGATGCAATGAGAACGTC 3388
 QY 841 ProSerSerSerGlyArgHisGlnGlyArgThrThrLysArgHisTyrArgLysSerVal 860
 DB CCATCTTCCAGTGAAGGCATGAAGAGAAAGAACTAACAAACGGCATTACCGAAATCTGTA 3448
 QY 861 ArgSerArgSerArgHisGlyTyrThrSerArgProLysLeuArgIleLeuAsnValSer 880
 DB AGAGTGTGCTCTGACATGAAAAAATCTCAAGCCCAAAATTTAGAAATTTGTAATGTTTCA 3508
 QY 881 AsnLysGlyAspArgValValGluCysGlnLeuGlnThrHisAsnArgLysMetValThr 900
 DB AATTAAGAGAACCGAGTGAAGTGAATGTCAATTGAGCTCATTAATGAGAAATAGCTTACA 3568
 QY 901 PheLysPheAspLeuAspGlyAspAsnProGluGluIleAlaThrIleMetValAsnAsn 920
 DB TTCAAATTTGACTAGTGTGACCAACCCGAGAGATGACCAATTAATGATGTAACAAT 3628
 QY 921 AspPheIleLeuAlaIleGluArgGluSerPheValAspGlnValArgGluIleIleGlu 940
 DB GACTTATTTCTACAAATGAGAGAGAGTGTGTGTGATCAAGTGCAGAAATTAATTTGAA 3668
 QY 941 LysAlaAspGluMetLeuSerGluAspValSerValGluProGluGluLysAspGlnGlyLeu 960
 DB AAGCTGATGAATGCTCAATGAGATGTGATGTGAAACAGAGGATGATCAGGATTTG 3748
 QY 961 GluSerLeuGlnGlyLysAspAspTyrGlyPheSerGlySerGlnLysLeuGlnGlyGlu 980
 DB GAGGCTTACAAAGAAAGATGACTATGCTTTCAAGTTCTCAGAAATTTGAAAGAGAG 3808
 QY 981 PheLysGlnProIleProAlaSerSerMetProGlnGlnIleGlyIleProThrSerSer 1000
 DB TTCAAAACCAACCAATCTGTGCTTCCATGCCACAGCAAAATAGGCATTTCTACACATGTTCT 3868
 QY 1001 LeuThrGlnValValHisSerAlaGlyArgArgPheIleValSerProValProGluSer 1020
 DB TTAACTCAAGTGTGCTTCTGCGGAGAGCGGTTTATATGATGCTCTGTGCGCAAGAGC 3928
 QY 1021 ArgLeuArgGluSerLysValPheProSerGluIleThrAspThrValAlaAlaSerThr 1040
 DB CGATTACGAATCAAAAGTTTCCCGAGTAATAACATACATACATGCTGCTCTTACAA 3988
 QY 1041 AlaGlnSerProGluMetLeuMetLysSerHisSerAlaSerSerLeuSerLeuGlnAla 1060
 DB GCTCAGAGCCCTGGAAGAAGTGTCTCAGCTGCAATCATCTTATGCTTACAAACAGGCC 4048
 QY 1061 PheSerGluLeuArgArgAlaGlnMetThrGluGlyProAsnThrAlaProProAsnPhe 1080
 DB TTTTCTCAACTTACAGCTGCCAAATACAGAGAGAACCAANACAGACCTCCAAACTTT 4108
 DB 1081 SerHisThrGlyProThrPheProValValProProPheLeuSerSerIleAlaGlyVal 1100

Db 4109 AGTCATACAGGACCAACATTTCCAGTACCTCTTTCTTAAGTAGCATTCGAGATC 4168
Qy 1101 ProthThrrAlaAlaAlaThrrAlaProValProAlaThrrSerSerProProAsnAsp11e 1120
Db 4169 CCAACVACAGACAGACAGCCAGACAGCTCCCTGGACACAGAGCCCTCTTAATAGACATT 4228
Qy 1121 SerThrrSerVal11eGlnSerGluValThrrValProthrrGluGluGly11AlaGlyVal 1140
Db 4229 TCCACATAGTAATTCAGTCTGAGGTACAGTGCCTGAAAGAGGGGATTCGCGAGTT 4288
Qy 1141 AlaThrrSerThrrGlyVal1ValThrrSerGlyGlyLeuPro11eProProValSerGluSer 1160
Db 4289 GCCACAGACAGAGGTGGTACTTCAGTGGTCTCCCATACACCTGTCTGAAATCA 4348
Qy 1161 ProValLeuSerSerValValSerSer11eThrr11eProAlaVal1Ser11eSerThr 1180
Db 4349 CCAGTACTTTCCACGTAAGTTTCAAGTATCACATACCTGCAAGTTGCTCAATCTACT 4408
Qy 1181 ThrrSerProSerLeuGlnValProThrrSerThrrSerGlu11Val1ValSerSerThrrAla 1200
Db 4409 ACATCCCGTCACTTCAAGTCCCAACATCCACATCAGATCGTTGTTCTAGTACAGCA 4468
Qy 1201 LeuThrrProSerValThrrValSerAlaThrrSerAlaSerAlaGlySerThrrAlaThr 1220
Db 4469 CTGATCTTCAGTAACAGTTTCAGCAATTCAGCTCTGCAAGGAGAGTACTGCTACC 4528
Qy 1221 ProGlyProLysProProAlaVal1ValSerGln11Ala11eGlySerThrrThrrVal1Gly 1240
Db 4529 CCAGGTCTTAAGCTCCAGCTGTAATCTCAGACAGCAGACAGGACACACTACTGAGGA 4588
Qy 1241 AlaThrrLeuThrrSerValSerThrrThrrThrrSerPheProSerThrrAlaSerGlnLeuSer 1260
Db 4589 GCACATTAACATCAGTTTCTACACACATTCATTCACAGACAGCTCTACAGCTGTSC 4648
Qy 1261 11eGlnLeuSerSerSerThrrSerThrrProThrrLeuAlaGluThrrVal1Val1SerAla 1280
Db 4649 ATTACGTTTACAGCACTACTTCTACTCTTACTTACTTAACACCGGAGTATGAGGCA 4708
Qy 1281 HisSerLeuAspLysThrrThrrSerHisSerSerThrrGlyLeuAlaPheSerLeuSerAla 1300
Db 4709 CACTCACTAATAGATGATCTCATAGCAGTACCACTGATTTGGTTTCTCTCTGCA 4768
Qy 1301 ProSerSerSerSerSerProGlyAlaGlyVal1SerSerTyrl1eSerGlnProGlyGly 1320
Db 4769 CCATCTCTCTCTCTCTCTGAGCAGCAGAGTGTCTAGTATATTTCTCAGCTGTGGG 4828
Qy 1321 LeuHisProLeuVal11eProSerVal11eAlaSerThrrPro11eLeuProGlnAla11a 1340
Db 4829 CTGCATCTTTGGTCACTCATAGATGATGCTTCTACTCTATTTCTTCCCAAGCAGCA 4888
Qy 1341 GlyProThrrSerThrrProLeuLeuProGlnValProSer11eProProLeuVal1GlnPro 1360
Db 4889 GGACCTACTTCTACACTTATTAACCCAGATAGTATCCACCCCTTGGTACAGCT 4948
Qy 1361 ValAlaAsnValProAlaVal1GlnGlnThrrLeu11eHisSerGlnProGlnProAlaLeu 1380
Db 4949 GTTGCCAAATGTGCTCTGTACGACAGACACTAATTCATAGTCAAGCTCAACACAGCTTG 5008
Qy 1381 LeuProAsnGlnProHisThrrHisGlyProGluVal1AspSerAspThrrGlnProLysAla 1400
Db 5009 CTTCACACACAGCCCATCTACTCTTCTCTGAAAGTATTCATACACACCAAGAGCT 5068
Qy 1401 ProGly11eAspAsp11eLysThrrLeuGluGluLysLeuArgSerLeuPheSerGluHis 1420
Db 5069 CCGGATATGATGATCACTAAAGACTTGAAGAAAGAACTGGGCTCTGTTCACTGAAAC 5128
Qy 1421 SerSerSerGlyAlaGlnHisAlaSerValSerLeuGluThrrSerLeuVal11eGluSer 1440
Db 5129 AGCTCATCTGAGAGTCAAGATGCTCTGCTCACTGAGAGCTCACTAGTCAATAGAGAGC 5188
Qy 1441 ThrrValThrrProGly11eProThrrAlaValAlaProSerTyrl1eLeuThrrSerThr 1460
Db 5189 ACTGTACACAGGATCCCACTACTGCTGTGGACCAAGCAAACTCTGACTTCTACC 5248

Qy 1461 ThrrSerThrrCysLeuProProThrrAsnLeuProLeuGlyThrrVal1AlaLeuProValThr 1480
Db 5249 ACAAGTACTTGTCTTACACACCAATTTTACACTAGAAACAGTTCCTTCCAGTTACA 5308
Qy 1481 ProValValThrrProGlyGlnValSerThrrProValSerThrrThrrThrrSerGlyValLys 1500
Db 5309 CCAGTGTCAACCTGGGCAAGTTTCTACCCCAATGACACTATCATCATGAGAGTGA 5368
Qy 1501 ProGlyThrrAlaProSerLysProProLeuThrrLysAlaProVal11eProVal1GlyThr 1520
Db 5369 CTTGAACTCTCTCTTCCAAAGCACTCTAACAAGAGCTCGGCTGCGCAGTGGTACT 5428
Qy 1521 GluLeuProAlaGlyThrrLeuProSerGluGlnLeuProPheProGlyProSerLeu 1540
Db 5429 GAATCTCCACAGGATCTTACCCAGCAGACAGTCCACCTTTTCCAGAGCTTCTCTA 5488
Qy 1541 ThrrGlnSerGlnProLeuGluAspLeuAspAlaGlnLeuArgThrrLeuSerPro 1560
Db 5489 ACCCAGTCCAGCAACCTTAGAGATCTTAGATCTTCAATTGAGAAAGACACTTAGTCA 5548
Qy 1561 Glu***11eThrrValThrrSerAlaVal1GlyProValSerMetAla11aProThrrAla11e 1580
Db 5549 GAATATATCACAGTACCTTCTGCGTTGCTGTCTGTCATGGCGGCTCYAACAGCAATC 5608
Qy 1581 ThrrGluAlaGlyThrrGlnProGlnLysGlyVal1SerGlnVal1LysGluGlyProValLeu 1600
Db 5609 ACRGAAGCAGAAACAGCTCCAGAAAGGCTTTTCTCAATCAAGAAAGGCTGTCTTA 5668
Qy 1601 AlaThrrSerSerGlyAlaGlyVal1PheLysMetGlyArgPheGlnValSerValAlaAla 1620
Db 5669 GCAACTAGTTCAAGAGCTGGTGTTTTAAAGATGGAGCAATTTAGGTTTCTGTTGCGCA 5728
Qy 1621 AspGlyAlaGlnLysGluGlyLysAsnLysSerGluAspAlaLysSerVal1HisPheGlu 1640
Db 5729 GACGGTCCCAAGAAAGGATAAATAAGTCAAGAAATGCCAAAGCTGTCTTATTTGAA 5788
Qy 1641 SerSerThrrSerGluSerSerVal1LeuSerSerSerSerProGluSerThrrLeuValLys 1660
Db 5789 TCCAGACCTCAGAGTCTCTCAGTCTTACAGTATGATCAAGATGATACCTTGTGTA 5848
Qy 1661 ProGluProAsnGly11eThrr11eProGly11eSerSerAspValProGluSerAlaHis 1680
Db 5849 CCAGACCCGAAATGATTAACATCCCTGGATCTTCTAGATGTCAGAGAGTGCCTAC 5908
Qy 1681 LysThrrThrrAlaSerGluAlaLysSerAspThrrGlyGlnProThrrLysVal1GlyArgPhe 1700
Db 5909 AAAACTACTGCTCAGAGCAAGATCAGACACTGGCAGCTTACCAAGTTGAGCGTTT 5968
Qy 1701 GlnValThrrThrrAlaAsnLysVal1GlyArgPheSerValSerTyrl1eThrrGluAspLys 1720
Db 5969 CAGGTACAACTACAGCAACAAAGTGGTCTTCTGTATCAAAACTGAGACAG 6028
Qy 1721 11eThrrAspThrrLysGlyGlyProVal1AlaSerProPheMetAspLeuGluGln 1740
Db 6029 ATCACTGACACAAAGAAAGAGGACAGTGGATCTCTCTTTATGATTTGAAACA 6088
Qy 1741 AlaValLeuProAlaVal11eProLysLysGluLysPProGluLeuSerGluProSerHis 1760
Db 6089 GCTGTCTTCTCTGTGTATCCAAAGAAAGAAAGAACCTGAACTGACAGCTTCAAT 6148
Qy 1761 LeuAsnGlyProSerSerAspProGluAlaAlaPheLeuSerArgAspVal1AspAspGly 1780
Db 6149 CTAAATGGCGCTCTTCTTACCCCGAGCGCTTTTAAAGTGGATGGAGTATGAT 6208
Qy 1781 SerGlySerProHisSerProHisGlnLeuSerSerLysSerLeuProSerGlnAsnLeu 1800
Db 6209 TCCGATAGTCCACATCGGCCCATCAGCTGAGTCAAAAGCTTCTTCCAGATCTA 6268
Qy 1801 SerGlnSerLeuSerAsnSerPheAsnSerSerTyrl1eMetSerSerAspAsnGluSerAsp 1820
Db 6269 AGTCAAAAGCTTATGATTAATTCATTTAACTCTCTTACATGATGATGACAAATAGTCAAT 6328

```

QY 1821 11eGluaSpGluAspLeuSLeuGluLeuArgArgLeuArgAspLeuSLeuSLeuGlu 1840
DB 6329 ATCGAAGATGAAACTTAAGTTAGAGCTGCGACGACTGAGATTAACATCTCAAGAG 6388
QY 1841 11eGluaSpGluInserArgGlnuSLeuGlu11eGluSLeuLeuYrThrLysLeuGly 1860
DB 6389 ATTCAGAGACTGCAAGTGGCCAGGAAGCAGAAATTAATCTTTGATATCAACATCGGCG 6448
QY 1861 LysValProProAlaVal11eLeuProProAlaAlaProLeuSergLysValGArgArg 1880
DB 6449 AAGGTGCCCCCTGCTGTTATATTCCTCCGAGCTGCTCCCTTTCAGGGAGAAACAGCA 6508
QY 1881 ProThrySergLysSergLysSergSergSergSergSergSergSergSergSergSerg 1900
DB 6509 CCACCTTAAGCAAGGACGACCAATCTAGTCAGAGAGCTTCCTGGGGAATTAAGCCCC 6568
QY 1901 GlnLeuSergLysLeuSergLysLeuSergLysLeuSergLysLeuSergLysLeuSerg 1920
DB 6569 CAGCTTTCAAGTAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6628
QY 1921 LeuHisProProGluVal11eProGluSergLysLeuSergLysLeuSergLysLeuSerg 1940
DB 6629 CTCACCTTCCTGCGCAACATCCAGAGTCCGCGCAGATCAGCTTACAGCCCTTAAG 6688
QY 1941 ProSergProSergSergSergSergSergSergSergSergSergSergSergSerg 1960
DB 6689 CCACTTCCTCCAGTACCACTCTTATTCAGCTTCACAGTATGATGCTGCTGCTGCTGCT 6748
QY 1961 ProSergLeuSergLysLeuSergLysLeuSergLysLeuSergLysLeuSergLysLeu 1980
DB 6749 CCAAGCTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6808
QY 1981 AsnSergLysLeuSergLysLeuSergLysLeuSergLysLeuSergLysLeuSergLys 2000
DB 6809 AACAGCCAAAGCCCGCCAGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6868
QY 2001 ThrAspAspLeuHisSLeuValAspAsnTrpAlaArgAspAlaMetAsnLeuSergLys 2020
DB 6869 ACAGATGACTTGCACAACTTGGTGAACATTTGGGCCGCAAGTGCATGATCTCTCAGGC 6928
QY 2021 ArgArgSergLysSergLysSergLysSergLysSergLysSergLysSergLysSergLys 2040
DB 6929 AGGAGAGAGAGCAAGGCGACATGATTAAGAGGGCCCTGGAATGGCAAGAGATCTCT 6988
QY 2041 AlaProGluGlnLeuSergLysSergLysSergLysSergLysSergLysSergLysSerg 2060
DB 6989 GCACCTGCGGCAACTGCTCATCTCCATGACCTCGAAGCTGGGTGCTGCTGCTGCTGCT 7048
QY 2061 AlaAlaSergLysSergLysSergLysSergLysSergLysSergLysSergLysSergLys 2080
DB 7049 GCACATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 7108
QY 2081 GlyPheProAlaThrProPheGlyAlaGlnTrpSergLysLeuSergLysLeuSergLys 2100
DB 7109 GGGCTTCCAGTACCCCATTTGGCGCTCAATGAGAGTGGAGCGGGTGGCCGACGACAG 7168
QY 2101 ProLeuGluGlnPheGlnProValGlyThrAlaSergLysLeuSergLysLeuSergLys 2120
DB 7169 CCACTTGGCGCACTTCCAACTGGGGAAGTGGCTCTTTCAGAAATTTCAACATCAGCAAT 7228
QY 2121 LeuGlnLysSergLysSergLysSergLysSergLysSergLysSergLysSergLysSerg 2136
DB 7229 TTGGCAAAATCATCAGCAACCCCGAGGCTTCAACTGCGGAGCAACT 7276

```

```

RESULT 10
PCT-US04-02652-454
; Sequence 454, Application PC/TUS0402652
; GENERAL INFORMATION:
; APPLICANT: CARCELIL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: RHEUMATOID ARTHRITIS, METHODS OF DETECTION AND USES THEREOF
; CURRENT APPLICATION NUMBER: PCT/US04/02652

```

```

; CURRENT FILING DATE: 2004-01-30
; NUMBER OF SEQ ID NOS: 50231
; SOFTWARE: FASTSEQ for windows Version 4.0
; SEQ ID NO 454
; LENGTH: 9639
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(9639)
; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1-7)
PCT-US04-02652-454

Alignment Scores:
Pred. No.: 0 Length: 9639
Score: 10567.50 Matches: 2096
Percent Similarity: 98.13% Conservative: 0
Best Local Similarity: 98.13% Mismatches: 11
Query Match: 97.74% Indels: 29
DB: Gaps: 1

US-10-010-720-14 (1-2136) x PCT-US04-02652-454 (1-9639)

QY 1 MetSergLysValAlaAlaGluSergLysSergLysSergLysSergLysSergLysSergLys 20
DB 956 ATGCTGCGGCGCGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1015
QY 21 ProAlaProAlaProLysAsnGlySergSergSergSergSergSergSergSergSergSerg 40
DB 1016 CCGGCTTCCTGCCCCCAAGATGGCTCAGCTCGATTCCTCGGGGAGAACTGGA 1075
QY 41 AlaAlaAlaAlaAspAlaValThrGlyArgThrGluLysArgArgArgArgArgArgArg 60
DB 1076 GCGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1135
QY 61 MetAspLysAspSergArgGlyAlaAlaAlaThrThrThrThrThrThrThrThrThrThr 80
DB 1136 ATGAGCAAGAGACCGCGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTTC 1195
QY 81 ArgArgSergVal11eCysAspSergAsnAlaThrAlaLeuGluLeuProGluLeuProLeu 100
DB 1196 CGCGGAGCGCTCATCTGNGACTCCCAATGCCATGCACTGAGCTTCCGGCTTCTCTT 1255
QY 101 SerLeuProGlnProSergLysProAlaAlaValProGlnSergLysProProProHis 120
DB 1256 TCCCTGCCAGCCGACGATCCCGCGGCTGCTCCGAGAGTGTCTCCACCGAGKCCCGAC 1315
QY 121 ArgGlnLysThrValThrAlaThrAlaThrSergLysValAlaGlnGlnProProAlaAla 140
DB 1316 CGGGAAGAGACCGTGAACCGCCACCGCACTTCCAGGTAGCCAGAGCCCTCCAGCGCT 1375
QY 141 AlaAlaProGluGlnGlnAlaValAlaGlyProAlaProSergLysThrValProSergSerg 160
DB 1376 RCCGCCCCGGGAAAGCGCGCTGCGCGGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTG 1435
QY 161 SerLysAspArgProValSergLysProSergLysValGlySergLysGlnGlnProProPro 180
DB 1436 AGCAAAACCGCCGAGTGTCCAGCTTGTGGGAGCAAAAGAGAGCGCGCGCG 1495
QY 181 AlaArgSergLysSergLysGlySergLysSergLysSergLysSergLysSergLysSerg 200
DB 1496 GCGAGAGTGGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1555
QY 201 GlnAspAsp11eGluLysLeuGluLysThrValAlaGlyMetSergAsnAspGlyArgPhe 220
DB 1556 CAGATGATATCGAAGGCTGAGACCAAGCGCTGGAAATGCTTAAGATGGCGGCTT 1615
QY 221 LeuLysPheAsp11eGluLysLeuGluLysThrValGlySergPheLysThrValYrLysGlyLeuAsp 240
DB 1616 CTCAAGTTTACATCGAAATCGGAGAGGCTCTTTAAAGGCTCTCAAAAGGCTGAGAC 1675
QY 241 ThrGluThrThrValGluValAlaTrpCysGluLeuGlnAspArgLysLeuThrLysSerg 260

```

Db 1676 ACTGAACACCGTGAAGTCCTGCTGATCGACGATCGAAAAATTACAAAGCTCT 1735
 QY 261 GUAAGGUAAGPheLeuGluGluAglMetLeuysGlyLeuGlnHisProAsnIle 280
 Db 1736 GAGAGGCGAGATTTAAAGAAAGAGCTGAATGTTAAAGAGCTTTCAGCATCCCAATATT 1795
 QY 281 ValArgPheTyraAspSerTrpGluSerThrValIysGlyLysGlyIleValLeuVal 300
 Db 1796 GTTAGATTATATGATTCCTGGGAATCCACAGTAAAGAAAGAAAGAGTGCATTGTTGGTG 1855
 QY 301 ThrGluLeuMetThrSerGlyThrLeuSerThrTyrlleuysAspPheLysValMetLys 320
 Db 1856 ACTGAACCTTAGAGCTCTGGAACCTTAAACGATATCGAAAGAGTTTAAAGTATGAAG 1915
 QY 321 IleLysValLeuAspSerTrpCysArgGlnIleLeuysGlyLeuGlnPheLeuHisThr 340
 Db 1916 ATCAAGTCTTAAGAGCTGGTGCGTCAGATCCTTAAAGTCTTCAGTTTCTTCATACT 1975
 QY 341 ArgThrProProIleIleHisArgAspLeuLysCysAspAsnIlePheIleThrGlyPro 360
 Db 1976 CGAATCCTACCTTACATTCACCGGATCTTAAATGTACAAACATCTTATCACCGGCCCT 2035
 QY 361 ThrGlySerValLysIleGlyAspLeuGlyLeuValaThrleuysArgAlaSerPheAla 380
 Db 2036 ACTGCTCACTCAAGATTCGAGACCTCGCTCTGGCAACCTCGAAAGCGGAGCTTCTTTGCC 2095
 QY 381 LysSerValIleGlyThrProGluPheMetAlaProGluMetTyrgluLysTyraSp 400
 Db 2096 AAGAGTGTAGTAGTACCCAGAGTTTCATGCGCCCTGAGATGTATGAGGAATATGAT 2155
 QY 401 GluSerValAspValTyraIlePheGlyMetCysMetLeuGluMetAlaThrSerGlyLys 420
 Db 2156 GAATCCGTACGTTATATGCTTTGGAGATGTGATGCTTGAATGCTTACATCTGATATAT 2215
 QY 421 ProTySerGluCysGlnAsnAlaIleGlnIleTyraArgValThrSerGlyValLys 440
 Db 2216 CTTACTCGAGAGTCCAAATGCTGCGACATCTTACCTGCGTGAACAGTGGGAGTGAAG 2275
 QY 441 ProAlaSerPheAspLysValAlaIleProGluValLysGluIleIleGluGlyCysIle 460
 Db 2276 CCAGCCAGTTTTCAGCAAGTAGCAATTCCTGAGGTAGGAATTTATGGAAGATGCTATA 2335
 QY 461 ArgGluAsnLysAspGluAlaGlyTyrSerIleLysAspLeuAsnHisAlaPhePheGln 480
 Db 2336 CGACAAACAAAGATGAAACATATTCATCAAAAGACCTTTGAAACATGCTTTTCCAA 2395
 QY 481 GluGluThrGlyValArgValGluLeuAlaGluLysAspGlyGluLysIleAlaIle 500
 Db 2396 GAGGAACAGAGATACGGTGAATTTAGCAAGAAAGATGATGAGAAAAAATAGCCATA 2455
 QY 501 LysLeuTrpLeuArgIleGluAspIleLysLysLeuLysGlyLysTyrlsAspAsnGlu 520
 Db 2456 AAATATATGCTACGATATTGAAGATATTAAAGAAATTAAAGGAATATACAAAGTATGA 2515
 QY 521 AlaIleGluPheSerPheAspLeuGluArgAspValProGluAspValAlaGlnGluMet 540
 Db 2516 GCTAATGAGTTTCTTTGATTATGAGAGAGATGCTCCAGAAAGATGTTGACAAAGAAATG 2575
 QY 541 ValGluSerGlyTyrlValCysGluGlyAspHisLysLeuThrMetAlaLysAlaIleLysAsp 560
 Db 2576 GTTAGATCTGGAGTATGCTGTGAAGGTGATCACAGACCATGGCTAAAGCTATCAAGAGC 2635
 QY 561 ArgValSerLeuIleLysArgLysArgGluGlnArgIleLeuValArgGluGlnGlu 580
 Db 2636 AGAGTATCATTAATTAAAGGAACAGAGACGCGCAGTTGGTACGGAGACGACAAAG 2695
 QY 581 LysLysLysGlnGluLysSerSerLeuLysGlnGlnValGluGlnSerSerAlaSerGln 600
 Db 2696 AAAAAAAGACAGAAAGACGACATCTCAACACAGAGTAGAACAAATCCAGTGTCTCCAG 2755
 QY 601 ThrGlyIleLysGlnLeuProSerAlaSerThrGlyIleProThrAlaSerThrThrSer 620
 Db 2756 AAGAGATCAAGAGAGTCCCTTCTGCTAGACACCGGATACCTTACTCTTACCACTTCA 2815

QY 621 AlaSerValSerThrGlnValGluProGluGluAspGlnHisGlnGlnLeu 640
 Db 2816 GCTTCAGTTTCTACACAGTAGAAGAACTGAAAGAACTGAGCGAGATCAACATCAACACTTA 2875
 QY 641 GlnTyrglnGlnProSerIleSerValLeuSerAspGlyThrValAspSerGlyGlnGly 660
 Db 2876 CAGTACAGCAACCCAGATATATCTGTATTCTGATGGAGCGGTGACAGTGTCAAGGA 2935
 QY 661 SerSerValPheThrGluSerArgValSerSerGlnGlnThrValSerTyrglySerGln 680
 Db 2936 TCCCTGCTCTTCATGAGATCTCGAGTACAGCAGCAACAGACGATTCATATATGTTCCAA 2995
 QY 681 HisGlnGlnAlaHisSerThrGlyThrValProGlyHisIleProSerThrValGlnAla 700
 Db 2996 CATGACAGGACATTTTACAGGACGACGTCAGGAGCATATVACCTTACTGTCACAGCA 3055
 QY 701 GlnSerGlnProHisGlyValTyrlProProSerSerValAlaGlnGlyGlnSerGlnGly 720
 Db 3056 CAGTCTCAGCCCATGAGGTATATCCACCTCAAGTGTG----- 3094
 QY 721 GlnProSerSerSerSerLeuThrGlyValSerSerSerGlnProIleGlnHisProGln 740
 Db 3094 ----- 3094
 QY 741 GlnGlnGlnGlyIleGlnGlnThrAlaProProGlnGlnThrValGlnTyrlSerLeuSer 760
 Db 3095 -----CAGGAATATACGACAGACAGCCCTCTCTCAACAGACAGAGTATTCATCTTCA 3148
 QY 761 GlnThrSerThrSerSerGluAlaThrThrAlaGlnProValSerGlnProGlnAlaPro 780
 Db 3149 CAGACATCAACCTCCAGTAGGCGCACACTAGCAGACAGCACTGATGTCARCTCAAGCTCCA 3208
 QY 781 GlnValLeuProGlnIleSerAlaGlyLysGlnSerThrGlnGlyValSerGlnValAla 800
 Db 3209 CAAGTCTGCTCAAGTATCAAGCTGAAAAACAAGATCACTGAGGAGTCTTCAGAGTTGCT 3268
 QY 801 ProAlaGluProValAlaValAlaGlnProGlnAlaThrGlnProThrThrLeuAlaSer 820
 Db 3269 CTGCGACAGCCAGTTGCTGATGACACAGCCCAAGCTACCCAGCGACACTTGGCTTCC 3328
 QY 821 SerValAspSerAlaHisSerAspValAlaSerGlyMetSerAspGlyAsnGluAsnVal 840
 Db 3329 TCTGTAGACAGTGCATTCAGATGTGCTTCAGATGATGATGATGCAATGAGACGTC 3388
 QY 841 ProSerSerSerGlyArgHisGluGlyArgThrThrLysArgHisTyrlArgLysSerVal 860
 Db 3389 CCATCTTCAAGTGAAGGATGAGAAAGAACTAACAAACGGCATTTAACGAAATCTGTA 3448
 QY 861 ArgSerArgSerArgHisGluLysThrSerArgProLysLeuAsnArgIleLeuAsnValSer 880
 Db 3449 AGAGTCTGCTCTGACATTAATAAACTTACGCCCCAAATTTAAAGATTTGAATGTTTCA 3508
 QY 881 AsnLysGlyAspArgValAlaGluCysGlnLeuGluThrHisAsnArgLysMetValThr 900
 Db 3509 AATAAGAGAGACCGAGTAGTAGATGTCAATTAGAGCATATATAGAAAAATGTTTACA 3568
 QY 901 PheLysPheAspLeuAspGlyAspAsnProGluGluIleAlaThrIleMetValAsnAsn 920
 Db 3569 TTCAAATTTGACCTATATGATGATCAACCCCGAGAGATAGCAACAATTATGTGAACAAT 3628
 QY 921 AspPheIleLeuAlaIleGluArgGluSerPheValAspGlnValArgGluIleIleGlu 940
 Db 3629 GACTTATATCTACAGATATAGAGAGAGTCTTTGTGATCAAGGTGGAATAATTATTGA 3688
 QY 941 LysAlaAspGluMetLeuSerGluAspValSerValGluProGluGlyAspGlnGlyLeu 960
 Db 3689 AAAAGCTGATGAATGCTCAGTAGAGATGTCAGTGTGAACCAAGAGGTGATCAGGATGG 3748
 QY 961 GluSerLeuGlnGlyLysAspAspTyrlGlyPheSerGlySerGlnLysLeuGluGlyGlu 980
 Db 3749 GAGAGTTCACAAAGAAAGATGATATGAGCTTTTTCAGGTTTTCAGAAATTTGAAGAGAG 3808

QY 141 AlaAlaProGluGluGlnAlaValAlaGlyProAlaProSerThrValProSerSerThr 160
Db 1376 RCGGCCCCCTGGGGAACAGGCCGTGCGGGCCCTCCCTCGACTGTCGCCAGAGTACC 1435
QY 161 SerLysAspArgProValSerGlnProSerLeuValGlySerLysGluGluProPro 180
Db 1436 AGCAAAACCGCCAGTGTCCAGCCTTAGCTTGGGGAGCAAAAGAGAGGCCGCCGCG 1495
QY 181 AlaArgSerGlySerGlyGlyGlySerAlaLysGluProGlnGluLysSerGln 200
Db 1496 GCGAGAACTGGCAGCGGGCGGCGAGCGCCAGAGGCCACAGAGAACGAGCCACAG 1555
QY 201 GlnAspAsp1LeGluGluLeuGluThrLysAlaValGlyMetSerAsnAspGlyArgPhe 220
Db 1556 CAGATGATATCGAAGAGCTGAGACCAAGCCCTGGGAATGTTAAAGATGCCGCTTT 1615
QY 221 LeuLysPheAsp1LeGlu1LeGlyArgGlySerPheLysThrValTyrLysGlyLeuAsp 240
Db 1616 CTCAAGTTTGACATCGAAATCGGACAGAGGCTCTTTAAAGACGCTCAAAAGTCTGGAC 1675
QY 241 ThrGluThrThrValGluValAla1ATPCysGluLeuGlnAspArgLysLeuThrLysSer 260
Db 1676 ACTGAAACCAACCGTGAAGTGTGCTGTGTAAGTCAAGATCAAAATTAACAAAGTCT 1735
QY 261 GluArgGlnArgPheLysGluGluAlaGluMetLeuLysGlyLeuGlnHisProAsn1Le 280
Db 1736 GAGAGCAGAGATTTAAAGAGAGCTGAATGTTAAAGTCTTCAGCATCCCAATTT 1795
QY 281 ValArgPheTyrAspSerTyrGluSerThrValLysGlyLysLysCys1LeuVal 300
Db 1796 GTTAGATTATTAATCTCTGGAATCCACAGTAAAGAAAGATCATGTTTGTG 1855
QY 301 ThrGluLeuMetThrSerGlyThrLeuLysThrTyrLeuLysArgPheLysValMetLys 320
Db 1856 ACTGAACTTATGACGCTGGAACCTTAACCGTATCTGAAGAGTTTAAAGTATGATAG 1915
QY 321 IleLysValLeuArgSerTyrCysArgGln1LeuLysGlyLeuGlnPheLeuHisThr 340
Db 1916 ATCAAGTTCTTAAGAGCTGGTGCCTCAAGTCTTAAAGTCTTCAGTTTCTTCATCT 1975
QY 341 ArgThrProPro1Le1LeHisArgAspLeuLysCysAspAsn1LePhe1LeThrGlyPro 360
Db 1976 CGAATCTCACCTTATCATCCGCGATCTTAATGTACAAACATCTTATCACCGGCTCT 2035
QY 361 ThrGlySerValLys1LeGlyAspLeuGlyLeuValaThrLeuLysArgAlaSerPheAla 380
Db 2036 ACTGGCTGAGTCAAGATTTGAGACCTGCTGCGAACCTCGAAGCGGCTCTTTGGCC 2095
QY 381 LysSerVal1LeGlyThrProGluPheMetAlaProGluMetTyrGluGluLysTyrAsp 400
Db 2096 AAGAGTGTAGTAGTACCCCAAGTTTCAATGGCCCTGAGATGTATGAGGAATATGAT 2155
QY 401 GluSerValAspValTyrAlaPheGlyMetCysMetLeuGluMetAlaThrSerGluTyr 420
Db 2156 GAATCCCTTACGTTATGCTTTTGGATGTGATGCTTGAGATGGCTTACATCTGATAT 2215
QY 421 ProTyrSerGluCysGlnAsnAla1aGln1LeTyrArgArgVal1ThrSerGlyValLys 440
Db 2216 CTTTACTCGAGTCCCAAAATGTGCGACATTTACCGTGGCGGACAGGGGGTGAAG 2275
QY 441 ProAlaSerPheAspLysValAla1a1eProGluValLysGlu1Le1LeGluGlyCys1Le 460
Db 2276 CCAAGCAGTTTGAACAAGTAGCAATCTCTGAAGTGAAGAAATTTATGAAGATGATA 2335
QY 461 ArgGlnAsnLysAspGluArgTyrSer1LeLysAspLeuLeuAsnHisAlaPhePheGln 480
Db 2336 CGCAAAACAAAGATGAAGATATTCATCAAGACCTTTTGAACCAATGCTTTTCCAA 2395
QY 481 GluGluThrGlyValArgValGluLeuAlaGluGluAspAspGlyGluLys1LeAla1Le 500
Db 2396 GAGGAAACAGAGAGTACGGGTGGAATTAGCAAGAAAGATGAGAAAAAATAGCCATA 2455
QY 501 LysLeuTyrPheLysArg1LeGluAsp1LeLysLysLeuLysGlyLysTyrLysAspAsnGlu 520

Db 2456 AAATTAAGCTACTACTATTGAAGATATTAAAGAAATTTAAAGGAAAAATCAAAAGATATGAA 2515
QY 521 Ala1LeGluPheSerPheAspLeuGluArgAspValProGluAspValAlaGlnGluMet 540
Db 2516 GCTATTGAGTTTCTTTTGAATTTTGAAGAGATGTCCAGAAAGATGTTGCAACAAGAAATG 2575
QY 541 ValGluSerGlyTyrValCysGluGlyAspHisLysThrMetAlaLysAla1LeLysAsp 560
Db 2576 GTAGAGTCTGAGTATGTCTGTGAAGGTGATCAAGAACCATGGCTTAAAGCTATCAAGAAC 2635
QY 561 ArgValSerLeu1LeLysArgLysArgGluGlnArgGlnLeuValArgGluGluGln 580
Db 2636 AGAGTATCATTAATTTAAGAGAAACGAGACAGGCGAGTTGTCACGAGAGACACAGAA 2695
QY 581 LysLysLysGlnGluGluSerSerLeuLysGlnGlnValGluGlnSerAlaSerGln 600
Db 2696 AAAAAAAGCAGAAAGAGACAGTCTCAACAGCAGGTAGAACATCCAGTCTTCCAG 2755
QY 601 ThrGly1LeLysGlnLeuProSerAlaSerThrGly1LeProThrAlaSerThrThrSer 620
Db 2756 ACAGAAATCAAGACGCTCCCTTGTGTAGCACCGCATACCTACTGCTTACCACTTCA 2815
QY 621 AlaSerValSerThrGlnValGluProGluGluProGluLysAspGlnHisGlnGlnLeu 640
Db 2816 GCTTCACTTCTTACACAAAGTAACTGAAAGAACCTGAAAGACCTGACATCAACATCAACTA 2875
QY 641 GlnTyrGlnGlnProSer1LeSerValLeuSerAspGlyThrValAspSerGlyGlnGly 660
Db 2876 CAGTACCAAGCAACCCAGTATATCTGTGTATCTGATGGACGGTGTGACAGTGTCAAGGGA 2935
QY 661 SerSerValPheThrGluSerArgValSerSerGlnGlnThrValSerTyrGlySerGln 680
Db 2936 TCTTCTCTTTCAYAGATCTCGATGAGGACCCAAAGCAAGCAAGTTTATATGATCCCAA 2995
QY 681 HisGluGlnAlaHisSerThrGlyThrValProGlyHis1LeProSerThrValGlnAla 700
Db 2996 CATGAACAGGCACATTTTACAGGACAGTCCAGGGCATATVACCTTCTACTGTCAGAGCA 3055
QY 701 GlnSerGlnProHisGlyValTyrProProSerSerValAlaGlnGlyGlnSerGlnGly 720
Db 3056 CAGTCTCAGCCCAATGGGTATATCCACCTCAAGTGTG----- 3094
QY 721 GlnProSerSerSerLeuThrGlyValSerSerSerGlnPro1LeGlnHisProGln 740
Db 3094 ----- 3094
QY 741 GlnGlnGlnGly1LeGlnGlnThrAlaProProGlnGlnThrValGlnTyrSerLeuSer 760
Db 3095 -----CAGGGAATPACAGACAGAGAGCCCTCCCTCAACAGACAGTGCAGTATTCATTCA 3148
QY 761 GlnThrSerThrSerSerGluAlaThrThrAlaGlnProValSerGlnProGlnAlaPro 780
Db 3149 CAGACATCAACCTCCAGTGGGCGACACTACAGCAGCAGCTGAGTCAACCTCAAGCTCCA 3208
QY 781 GlnValLeuProGlnValSerAlaGlyLysGlnSerThrGlnGlyValSerGlnValAla 800
Db 3209 CAACTCTTGCCTCAAGTATAGCTGGAACAGAGTACTCAGGAGAGTCTTCAGGTGCT 3268
QY 801 ProAlaGluProValAlaValAlaGlnProGlnAlaThrGlnProThrThrLeuAlaSer 820
Db 3269 CTGCAAGAGCCAAATTGAGTAGACACAGMCCCAAGCTATCCAGCGCAACATTTGGCTTCC 3328
QY 821 SerValAspSerAlaHisSerAspValAlaSerGlyMetSerAspGlyAsnGluAsnVal 840
Db 3329 TCTGTAGACAGTGCACATTCAGATGTTGCTTCAGTATGATGCAATGCAATGCAAGCTC 3388
QY 841 ProSerSerSerGlyArgHisGluGlyArgThrThrLysArgHisTyrArgLysSerVal 860
Db 3389 CCAITCTTCCAGTGAAGGCAATGAAGAGAACTTACAAAGGCAATTTACCAAAATCTCTGA 3448
QY 861 ArgSerArgSerArgHisGluLysThrSerArgProLysLeuArg1LeLeuAsnValSer 880

Db 3449 AGGAGTGCCTCGACATGAAAAAATTCAAGCCCAAAATTTAGAAATTTGTAATGTTTCA 3508
QY 881 AsnLysGlyAspArgValValGluCysGlnLeuGluThrHisAsnArgLysMetValThr 900
Db 3509 AATTAAGAGACCGAGTAGTAGAATGTCAATTAGAGCTCATATATAGAAAAATGTTTACA 3568
QY 901 PheLysPheAspLeuAspGlyAspAsnProGluGluLeuAlaThrLleMetValAsnAsn 920
Db 3569 TTCAAAATTTGACCTAGATGTGACAAACCCGAGAGATAGCAAAATTAATGAGCAAT 3628
QY 921 AspPheLleLeuAlaLleGluArgGluSerPheValAspGlnValArgGluLleLeu 940
Db 3629 GACTTATTTTACCAATAGAGAGAGAGTCTTTGTGATCAAGTGCAGAAATTAATTTGAA 3688
QY 941 LysAlaAspGluMetLeuSerGluAspValSerValGluProGluGluAspGlnGlyLeu 960
Db 3689 AAGAGTGATGAATGTGCTCAGTAGAGATGTGAGTGTGAAACCAAGGTTGATCAGGATTTG 3748
QY 961 GluSerLeuGlnGlyLysAspAspTyrGlyPheSerGlySerGlnLysLeuGluGly 980
Db 3749 GAGAGTCTACAAAGAAAGGATGACTATGCTTTTCAGTTCTCAGAAATTTGAAAGAGAG 3808
QY 981 PheLysGlnProLleProAlaSerSerMetProGluGlnLleGlyLleProThrSerSer 1000
Db 3809 TTCAAAACAACTTCTGCTGCTCTCCATGCAACAGCAATAGGCATTTCTTACCTGCTCT 3868
QY 1001 LeuThrGlnValValHisSerAlaGlyArgArgPheLleValSerProValProGluSer 1020
Db 3869 TTAACTCAAGTTTTCATTTCTGCGGAGAGCGGTTTATAGTAGTCTGTGCGCAAGAAC 3928
QY 1021 ArgLeuArgGluSerLysValPheProSerGluLleThrAspThrValAlaAlaSerThr 1040
Db 3929 CGATTTCGAGAAATCAAAAGTTTTCCTCCAGTGAATTAACAGATTCAGTTGCTGCTCAAC 3988
QY 1041 AlaGlnSerProGlyMetLeuSerSerHisSerAlaSerSerSerLeuGlnGlnAla 1060
Db 3989 GCTCAAGCCCTGAAATGAACTGTCTCACTGCAATCATCTTATGCTTCAACAGGCC 4048
QY 1061 PheSerGluLeuArgArgAlaGlnMetThrGluGlyProAsnThrAlaProProAsnPhe 1080
Db 4049 TTTTTCGAAGTTTAGAGTGCCTCAATGACGAAGAGACCAAAACAGCAGCTTCCAACTTT 4108
QY 1081 SerHisThrGlyProThrPheProValValProProPheLeuSerSerLleAlaGlyVal 1100
Db 4109 AGTCATAACGAGACCAATTTCCAGTAGTACTCTTTCTTAAGTAGCAATGCTGAGAGTC 4168
QY 1101 ProThrThrAlaAlaThrAlaProValProAlaThrSerSerProProAsnAspLle 1120
Db 4169 CCAACTACAGACAGACAGACAGCAACAGTCTCTGCAACAGAGCCCTCTTAATGACAT 4228
QY 1121 SerThrSerValLleGlnSerGluValThrValProThrGluGluGlyLleAlaGlyVal 1140
Db 4229 TCCACTACAGTATTCAGTCTGAGTTTACGTGACGTCCTCAAGAGGGATGCTGAGATT 4288
QY 1141 AlaThrSerThrGlyValValThrSerGlyGlyLeuProLleProProValSerGluSer 1160
Db 4289 GCCACACAGACAGAGTGTGTAACCTCAGAGTGTCTCCCATACCACTGTGTGATCA 4348
QY 1161 ProValLeuSerSerValValSerSerLleThrLleProAlaValValSerLleSerThr 1180
Db 4349 CAGATACCTTCCAGCGTAGTTTCAAGATACCAATACCTGCAAGTGTCTCAAAATATCTACT 4408
QY 1181 ThrSerProSerLeuGlnValProThrSerThrSerGluLleValValSerSerThrAla 1200
Db 4409 AACTCCCGTCACTTCAAGTCCCAATCCACATCTGAGATGCTGTGTTCTAGTTACAGCA 4468
QY 1201 LeuTyrProSerValThrValSerAlaThrSerAlaSerAlaGlyGlySerThrAlaThr 1220
Db 4469 CTGTATCTCTTCAGTATCAGTTTCAGCACTTCAGCTTCGAGGGGAGTACTGCTACAC 4528
QY 1221 ProGlyProLysProProAlaValValSerGlnGlnAlaAlaGlySerThrThrValGly 1240
Db 4529 CCAGGTCTTAAGCTTCAAGCTGTAGTATCTCAGACAGACAGACAGCACTACTGTGGCA 4588

QY 1241 AlaThrLeuThrSerValSerThrThrThrSerPheProSerThrAlaSerGlnLeuSer 1260
Db 4589 GCCACTTAACATCAATTTTCTTCCACCACTTCACTTCCCAAGACAGCTTACAGCTGTSC 4648
QY 1261 LleGlnLeuSerSerSerThrSerThrProThrLleuAlaGluThrValValSerAla 1280
Db 4649 ATTCAGCTTAGCAGCAGTACTTCTTCACTTCTTACTTGTAGTGAACCGGTGATGAGCGCA 4708
QY 1281 HisSerLeuAspLysThrSerHisSerSerThrThrGlyLeuAlaPheSerLeuSerAla 1300
Db 4709 CACTCACTGATTAAGACATCTCATACAGTACAACTGATGAGTGGCTTCTCTCTCTGCA 4768
QY 1301 ProSerSerSerSerSerProGlyAlaGlyValSerSerTyrLleSerGlnProGlyGly 1320
Db 4769 CCACTTCTCTTCTCTCTCTCTGAGAGAGAGTCTAGTTATTTCTCAGCTGTGAGG 4828
QY 1321 LeuHisProLeuValLleProSerValLleAlaSerThrProLleLeuProGlnAlaAla 1340
Db 4829 CTGCATCTTTGGTCAATTCATCAGTAGATGTTTACTTCTTCTTCTTCTTCTTCTTCTTCT 4888
QY 1341 GlyProThrSerThrProLeuLeuProGlnValProSerLleProProLeuValGlnPro 1360
Db 4889 GGACTTACTTCTTACACCTTATTTATCCCAAGTACCTAGATCCCACTTGTGACAGCCT 4948
QY 1361 ValAlaAsnValProAlaValGlnGlnThrLleLleHisSerGlnProGlnProAlaLeu 1380
Db 4949 GTTGCCAAATGTGCTGCTGTGACGACGACATTAATTAATGATGAGCTCAACAGCTTTG 5008
QY 1381 LeuProAsnGlnProHisThrHisCysProGluValAlaAspSerAspThrGlnProLysAla 1400
Db 5009 CTTCACCAACAGGCCCATCACTCATTTGCTGAAGTATGATGATACAAACCAAGCT 5068
QY 1401 ProGlyLysAspAspLleLysThrThrLeuGluGluLysLeuArgSerLeuPheSerGluHis 1420
Db 5069 CCTGGAATGATGACATTAAGACTTGAAGAAAGAACTCGGCTGTGTCAGTGAACAC 5128
QY 1421 SerSerSerGlyValAlaGlnHisAlaSerValSerLeuGluThrSerLeuValLleGluSer 1440
Db 5129 AGCTACTGAGACCTGACATGCTGCTGTCTCACTGAGAACCTCACTATGATAGAGAGC 5188
QY 1441 ThrValThrProGlyLleProThrThrAlaValAlaProSerLysLeuLeuThrSerThr 1460
Db 5189 ACTGTACACAGAGCATCCCACTACTGCTGTGACCAAGCAAACTCTGCACTTCTAC 5248
QY 1461 ThrSerThrCysLeuProProThrAsnLeuProLeuGlyThrValAlaLeuProValThr 1480
Db 5249 ACAAGTACTTGTCTTACCAACCAATTTTACCACTAGGAACAGTTGCTTGGCCAGTTTACA 5308
QY 1481 ProValValThrProGlyGlnValSerThrProValSerThrThrSerGlyValLys 1500
Db 5309 CCAAGTGTACACCTGAGGCAAGTTTCTTACCAGATGACATCTACATCAGAGAGTGA 5368
QY 1501 ProGlyThrAlaProSerLysProProLeuThrLysAlaProValLeuProValGlyThr 1520
Db 5369 CCTGGAACCTGCTCCCTCCCAAGCACCTTAACTTAAGGCTCCGAGTGTGCAAGTGTACT 5428
QY 1521 GluLeuProAlaGlyThrLeuProSerGluGlnLeuProProPheProGlyProSerLeu 1540
Db 5429 GAACTTCCAGCAGGATCTTACCAACGAGCAGAGCTGCACTTTTCCAGAGCACTTCTCTCA 5488
QY 1541 ThrGlnSerGlnGlnProLeuGluAspLeuAspAlaGlnLeuArgThrThrLeuSerPro 1560
Db 5489 ACCGATCCCAACCACTTACAGATCTTGTATGCTCAATTTAGAGAAACATTAAGTCCA 5548
QY 1561 Glu***LleThrValThrSerAlaValGlyProValSerMetAlaAlaProThrAlaLle 1580
Db 5549 GAGATATATCAGATGATCTTCTGCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5608
QY 1581 ThrGluAlaGlyThrGlnProGlnLysGlyValLysGlnValLysGlnGlyProValLeu 1600
Db 5609 ACRGAGACGAGAACAGCTCAGAGAGGAGTGTCTTCAAGTCAAAAGAGGCTGTGCTCA 5668

Db 1016 CCGGCTCTCTGCCCCCAAGAAATGCTCCAGCTCCGATTCCTCCGTGGGGGAGAAACTGGGA 1075

Qy 41 AAlaAlaAlaAlaAlaAlaValThrGlyArgThrGluGluTyrArgArgArgGlnSerThr 60

Db 1076 GCCCGCGCGCCCGCAGCTGTGACCGGAGAGACGAGAGATACAGCGCGCCCGCCGACACT 1135

Qy 61 MetAspLysAspSerArgGlyAlaAlaAlaThrThrThrThrThrGluHisArgPhePhe 80

Db 1136 ATGACAAAGAGACGCCGTGGGGGGCGGCGACCACTACCACTGAGCACCGCTTCTTCC 1195

Qy 81 ArgArgSerValIleCysAspSerAspAlaThrAlaLeuGluLeuProGluLeuProLeu 100

Db 1196 CGCGGAGGCTCATCTGAGACTCCATGACCTGACCTGAGAGCTTCCGGGCTTCTTCT 1255

Qy 101 SerLeuProGluProSerIleProAlaAlaValProGluSerAlaProProGluProHis 120

Db 1256 TCCCTGGCCCGACCGACATCCCGGGGCTGTCCCGAGAGTGCTCCACCGGACCCCCAC 1315

Qy 121 ArgGluGluThrValThrAlaThrAlaThrSerGluValAlaGluGluProProAlaAla 140

Db 1316 CCGGAAAGAGACCGTGACCGGCGACCGCACTCCAGGTAGCCAGCGCTCCACGCGCT 1375

Qy 141 AlaAlaProGluGluGluAlaValAlaGlyProAlaProSerThrValProSerSerThr 160

Db 1376 RCCGCCCCCTGGGAAACAGGCCCTCGCGGGCCCTGCCCCCTCGACTGTCCACGAGTACC 1435

Qy 161 SerLysAspArgProValSerGluProSerLeuValGlySerLysGluGluProProPro 180

Db 1436 AGCAAAAGACCGCCAGGTGCCAGCTTGTGGGGAGCAAAAGAGGAGCGCGCGCG 1495

Qy 181 AlaArgSerGlySerGlyGlySerAlaLysGluProGluGluLysArgSerGluGln 200

Db 1496 GCGAGAGGTGGCGCGCGCGGCGACCGCAAGAGAGCAACGAGAGAACCGGACCGAG 1555

Qy 201 GlnAspAspIleGluGluLeuGluThrLysAlaValGlyMetSerAspAspGlyArgPhe 220

Db 1556 CAGAGATGATTCGAAGAGCTGGAGACCAAGCCCGTGGAATGCTAACGATGCGCTT 1615

Qy 221 LeuLysPheAspIleGluIleGlyArgGlySerPheLysThrValTyrLysGlyLeuAsp 240

Db 1616 CTCAGATTGACATCGAAATCGGACAGAGCTCTCTTAAAGCGTCTACAAAGCTCGGAC 1675

Qy 241 ThrGluThrThrValGluValAlaATP-CysGluLeuGlnAspArgLysLeuThrLysSer 260

Db 1676 ACTGAACCAACCGTGGAAGTGGCTGTGTAACCTCAGAGATCAAAATTAACAAGTCT 1735

Qy 261 GlnArgGluArgPheLysGluGluAlaGluMetLeuLysGlyLeuGlnHisProAlaIle 280

Db 1736 GAGAGCGACGATTTAAAGAAAGCTGAATGTAAAGGTCTTACAGCATCCCAATATT 1795

Qy 281 ValArgPheLysAspSerTrpGluSerThrValLysGlyLysLysCysIleValLeuVal 300

Db 1796 GTTAGATTATATCATCTCTGGGATCCACGATTAAGTAAGAAAGAGTGCATTGTTGGTG 1855

Qy 301 ThrGluLeuMetThrSerGlyThrLeuLysThrTyrLeuLysArgPheLysValMetLys 320

Db 1856 ACTGAACCTATGACGTCTGGAACACTTAAACGTATCTGAAGAGGTTTAAAGTATGAAG 1915

Qy 321 IleLysValLeuArgSerTrpCysArgGlnIleLeuLysGlyLeuGlnPheLeuHisThr 340

Db 1916 ATCAAAAGTTCTAAAGAGCTGGTGGCGGTCAAGATCTTAAAGGTCTTCAAGTTCTTCA 1975

Qy 341 ArgThrProProIleIleHisArgAspLeuLysCysAspAsnIlePheIleThrGlyPro 360

Db 1976 CGAAGCTCCACCTATCATTCACCGCGATCTTAAATGTGACAACTTTATACCGGCGCT 2035

Qy 361 ThrGlySerValLysIleGlyAspLeuGlyLeuAlaThrLeuLysArgLysSerPheAla 380

Db 2036 ACTGGGCTCAAGTGAAGTGAAGCTCGGTCTGGCAACCTCGAAGCGGGCTTCTTTGGC 2095

Qy 381 LysSerValIleGlyThrProGluPheMetAlaProGluMetThrGluLysTyrZnAp 400

Db 2096 AAGAGTGTATAGTACCCAGAGTTCATGGCCCCCTGAGATGTATAGAGAAATATATGAT 2155

Qy 401 GluSerValAspValTyrAlaPheGlyMetCysMetLeuGluMetAlaThrSerGluTyr 420

Db 2156 GAATCGGTTCACCTTATATGCTTTGGGATGTGATGTGATGAGTGCATCATCTGAATAT 2215

Qy 421 ProTyrSerGluCysGlnAsnAlaAlaGlnIleTyrArgArgValThrSerGlyValLys 440

Db 2216 CTTTACTCGGAGTGCAGAAATGCTGCRAGATCATCGTGGCTGACCAAGTGGGGTGAAG 2275

Qy 441 ProAlaSerPheAspLysValAlaIleProGluValLysGluIleIleGluGlyCysIle 460

Db 2276 CAGCAAGTTTGAACAAGTAGCAATTCGAAAGTGAAGAAATTTATGAAGATGATATA 2335

Qy 461 ArgGlnAsnLysAspGluArgTyrSerIleLysAspLeuLeuAsnIleAlaPhePheGln 480

Db 2336 CGACAAACAAATATGAAGATATTCATCACTAAAGACTTTTGAACCATGGCTTTTCCAA 2395

Qy 481 GluGluThrGlyValArgValGluLeuAlaGluLysAspGlyLysIleAlaIle 500

Db 2396 GAGGAAACAGAGTACGGGTAGAAATTAGCAGAGAGAGATGAGAGAAATAATAGCATTA 2455

Qy 501 LysLeuThrLeuAspGluIleGluAspIleLysLysLeuLysGlyTyrTyrLysAspAsnGlu 520

Db 2456 AATTTATGGCTACGTATTAAGATATTAAAGAAATTAAGGAAATTAACAAAGTAAATGAA 2515

Qy 521 AlaIleGluPheSerPheAspLeuGluArgAspValProGluAspValAlaGluGluMet 540

Db 2516 GCTATTGAGTTTCTTTATATTTAGAGAGAGATGTCCCAAAATGTGTGACAAAGAAATG 2575

Qy 541 ValGluSerGlyTyrValCysGluGlyAspHisLysThrMetAlaLysAlaIleLysAsp 560

Db 2576 GTAGAGTCTGGGTATGTCTGTGAAGGTGATCAACAAGCATGAGCTTAAGCTATCAAGAC 2635

Qy 561 ArgValSerLeuIleLysArgLysArgGluGlnArgGlnLeuValArgGluGluGlnGlu 580

Db 2636 AAGATATCTTAAATTAAAGAGAAACGAGAGCGGAGCTGTGACGGGAGGAGCAAGAA 2695

Qy 581 LysLysLysGluGluGluSerSerLeuLysGlnGlnValGluGlnSerSerAlaSerGln 600

Db 2696 AAAAAAACAAGAAAGAGACGACTCCAAACAGCAGTGAACATTCAGTCTTCCAG 2755

Qy 601 ThrGlyIleLysGluLeuProSerAlaSerThrGlyIleProThrAlaSerThrThrSer 620

Db 2756 ACAGGAATCAACAGAGCTCCTCTGTGTACACCGGATGACTCTTCTTACCACTTCA 2815

Qy 621 AlaSerValSerThrGlnValGluProGluGluProGluAlaAspGlnHisGlnGlnLeu 640

Db 2816 GCTTCAGTTTCTACACAAAGTAGAAACCTGAAGAACTGAGGCAAGTCAACATCAACAAC 2875

Qy 641 GlnTyrGlnGluProSerIleSerValLeuSerAspGlyThrValAspSerGlyGlnGly 660

Db 2876 CAGTACAGCAACCCAGATATCTGTGTATCTGATGAGAGCGTTTACAGTGTACAGGA 2935

Qy 661 SerSerValPheThrGluSerArgValSerSerGlnGlnThrValSerTyrGlySerGln 680

Db 2936 TCCCTGTCTTCAVAGAAATCTGAGTAGCAGCACAACAGATTCAATATGTTCCCA 2995

Qy 681 HisGluGlnAlaHisSerThrGlyThrValProGlyHisIleProSerThrValGlnAla 700

Db 2996 CATGAACAGCACATTTCAAGGACAGCAGATCCAGGGCAATATCTTCTGCTCCAGAGA 3055

Qy 701 GlnSerGlnProHisGlyValTyrProProSerSerValAlaGlnGlyGlnSerGlnGly 720

Db 3056 CAGTCTCAGCCCAATGGGTATATTCACCTCAAGTGTGGACAGGGGCGAGGCCAGGT 3115

Qy 721 GlnProSerSerSerSerLeuThrGlyValSerSerSerGlnProIleGlnHisProGln 740

Db 3116 CAGCATCTCTCAAGTAGCTTAACAGGGGTTTATCTTCCCAACCATACAAATCTCTCAG 3175

Qy 741 GlnGlnGlnGlyIleGlnGlnThrAlaProProGlnGlnThrValGlnTyrSerLeuSer 760

Db 3176 CAGCAGCAGGAGATACAGCAGACAGCCCTCTCTCAACAGACAGTGCATTAATCACTTCA 3235

QY 761 GlnThrSerThrSerSerGluAlaThrThrAlaGlnProValSerGlnProGlnAlaPro 780
DB 3236 CAGACATCAACCTCCAGTAGAGGCCACTACTACGACACGCCAGTAGTCARCTCAAGCTCCA 3295
QY 781 GlnValLeuProGlnValSerAlaGlyLysGln----- 791
DB 3296 CAAGTCTTGCTCAAGATATACGTGGAAAAACAGCTTCCAGTTCCAGGACGATACCAACT 3355
QY 791 ----- 791
DB 3356 ATCCAAAGCGAACCTCGATGCCAGTTCGACACAACCCCTGGTTGTTCCAGTCCACTCT 3415
QY 791 ----- 791
DB 3416 GGTGCTCATTTCTCTTCAGTGGAGACAGCCGCTCCCTACTCCCTTCTCCTCAGTACCT 3475
QY 791 ----- 791
DB 3476 GTCTCATATTCCTCATATCATCCGATGTGTACGGCTCAGACAGGTTCTCATCC 3535
QY 791 ----- 791
DB 3536 CTTCCTCATCAATGCGACGTGCAGTATCTCAGCTCTGCTCAGTTCGCTCATCTGCT 3595
QY 791 ----- 791
DB 3596 ACACACGCTGCCATCCCGGGGGATATCACTGTGTTCTTAGTACGTTCCACCCCTTCTG 3655
QY 791 ----- 791
DB 3656 CAGCTGTGACTCAGGTCGCCAAGTCAAGTTCACCCAAGCTCTCAACACGAGAGTTGAG 3715
QY 791 ----- 791
DB 3716 TCCATGGGAATACAGCTAATCTTGGACAGCTGTGAGTTCCACTTTCCTCTGGAGAT 3775
QY 791 ----- 791
DB 3776 GTTCTGTACAGGGCTTCCCACTCGACTGCCACACAGTACCCAGAGATTCAAATATT 3835
QY 791 ----- 791
DB 3836 GCTCCCTCTTCAACGAGGCTTGTGTTGATCATCTTCAAGTCCATCCCTCCCATG 3895
QY 791 ----- 791
DB 3896 CCGACAGAACTACTGGCTACACCTGGAGTACTTCCACAGTGGTGAGACCTTATGGAA 3955
QY 791 ----- 791
DB 3956 TCAAACTTTTACTTCTTATGGGTGTGTAGAGAGCAGTTCAAGTCCCAAGCAGGA 4015
QY 792 ----- 792
DB 4016 GGGAGTTTACACAGCCGCCACATCACTCCACAGAACAGTTTGGAGAGTACTCAG 4075
QY 795 GlyValSerGlnValAlaProAlaGlnProValAlaValAlaGlnProGlnAlaThrGln 814
DB 4076 GGAAGTCTTCAAGGTGCTCTGTGACAGCCAGTTCAGTACACAGMCCCAAGCTACCCAG 4135
QY 815 ProThrThrLeuAlaSerSerValAspSerAlaHisSerAspValAlaSerGlyMetSer 834
DB 4136 CCGACCACTTTGGCTTCTCTGTAGACAGTGCACATTCAGATGTGCTTACAGATAGAGT 4195
QY 835 AspGlyAsnGluAsnValProSerSerSerGlyArgHisGlnGlyArgThrThrLysArg 854
DB 4196 GATGGCAATAGAACGTCATCTTCCAGTGAAGCATGAAGAACTAACAAAAACGG 4255
QY 855 HisTyrArgLysSerValAlaArgSerArgSerArgHisGlnLysThrSerArgProLysLeu 874
DB 4256 CATTAACGAAAAATCTGTAAAGAGTCGCTTCGACATGAAAAAACTTCAACCCCAAAATTA 4315
QY 875 ArgIleLeuAsnValSerAsnLysGlyAspArgValValGluCysGlnLeuGluThrHis 894

DB 4316 AGAATTTTAAATGTTTCAATTAAGAGACCCAGTGTGTGAATGTCAATTAGAGCTCAT 4375
QY 895 AsnArgLysMetValThrPheLysPheAspLeuAspGlyAspAsnProGlnGluIleAla 914
DB 4376 AATAGAAAATGGTTCATTTCAATTAATTTGACTAGATGTGACACACCCAGAGAGATAGCA 4435
QY 915 ThrIleMetValAspAsnAspPheIleLeuAlaIleGluArgGluSerPheValAspGln 934
DB 4436 ACAATATGTGTGAACATACCTTATTTAGCAATAGAGAGAGATGCTTTGTGATCAAA 4495
QY 935 ValArgGluIleIleGluLysAlaAspGluMetLeuSerGluAspValSerValGluPro 954
DB 4496 GTCCGAAAATTTATGAAAAAGCTGATGAATTCCTCAGTAGAGATTCAGTGGAAACCA 4555
QY 955 GluGlyAspGlnGlyLeuGlnSerLeuGlnGlyLysAspAspTyrGlyPheSerGlySer 974
DB 4556 GAGGGTGATCAGGGATTGGAGAGTCTACAAAGAAAGATGACTATGGCTTTTCAGGGTCT 4615
QY 975 GlnLysLeuGlnGluGlnPheLysGlnProIleProAlaSerSerMetProGlnGlnIle 994
DB 4616 CAGAAATGGAAAGAGAGTTCAAACACCAATTCGCGCTTCCATGCGACAGCAAAATA 4675
QY 995 GlyIleProThrSerSerLeuThrGlnValValHisSerAlaGlyArgArgPheIleVal 1014
DB 4676 GGCATTCCTACCAAGTCTTTAATCAAGTGTTCATTCGCGGGAAGCGGTTATATGTC 4735
QY 1015 SerProValProGluSerArgLeuArgGluSerLysValPheProSerGluIleThrAsp 1034
DB 4736 AGTCCGTGCCAGAAAGCCGATTAACAGATCAAAAGTTTCCCGAGTAAATMAACAGAT 4795
QY 1035 ThrValAlaAlaSerThrAlaGlnSerProGluLysAsnLeuSerHisSerAlaSerSer 1054
DB 4796 ACAGTGTGCTCTCAAGCTCAGAGCCCTGGAGTAACCTGTCTCACTGTGATCATCC 4855
QY 1055 LeuSerLeuGlnGlnAlaPheSerGluLeuArgArgAlaGlnMetThrGlnGlyProAsn 1074
DB 4856 CTTAGTCTCAACAGAGCCCTTTCTGAACCTTAGACGTGCCAATGACAGAGACCCAAAN 4915
QY 1075 ThrAlaProProAsnPheSerHisThrGlyProThrPheProValAlaProPheLeu 1094
DB 4916 ACAGCACTTCCAAACTTTAGTCAACAGACCAACATTTCCAGTATACCTCTTCTTCTTA 4975
QY 1095 SerSerIleAlaGlyValAlaProThrThrAlaAlaAlaThrAlaProValProAlaThrSer 1114
DB 4976 AGTAGATGTCTGAGATTCACCAACACAGACAGCCACAGCACTCCTGCAACAGC 5035
QY 1115 SerProProAsnAspIleSerThrSerValIleGlnSerGluValThrValProThrGln 1134
DB 5036 AGCCCTCTATATGACATTTCCATCATCAGTATTCAGTCTGAGGTTACAGTGCCACGTGA 5095
QY 1135 GlnGlyIleAlaGlyValAlaThrSerThrGlyValValThrSerGlyLeuProIle 1154
DB 5096 GAGGGGATGTCTGGAGTTGCCACACAGACAGGAGTGTAACTTCAGGTGCTCTCCCATTA 5155
QY 1155 ProProValSerGlnSerProValLeuSerSerValValSerSerIleThrIleProAla 1174
DB 5156 CCACTGTGTGTAATCACACAGTACTTCCAGCGTATTTCAAGTATTCACAAATACCTGCA 5215
QY 1175 ValValSerIleSerThrThrSerProSerLeuGlnValProThrSerThrSerGluIle 1194
DB 5216 GTTGTCTCAATATCTATCATATCCCGTCACTTCAAGTCCCAATCCACATCGAGATC 5275
QY 1195 ValValSerSerThrAlaLeuTyrProSerValThrValSerAlaThrSerAlaSerAla 1214
DB 5276 GTTGTCTTCAATGACACACAGTATCTCTCAGTAAAGTTCACAACTTCAGGCTCTGCA 5335
QY 1215 GlyGlySerThrAlaThrProGlyLysProLysProProAlaValValSerGlnGlnAlaIle 1234
DB 5336 GGGGGGAGTACTCTACCCAGGTTCTTAACTTCCAGCTCTAATATCTCAGCGGAGGCA 5395
QY 1235 GlySerThrThrValGlyAlaThrLeuThrSerValSerThrThrThrSerPheProSer 1254

Db 5396 GGCAGCACTACTGAGGAGCCATTAACATCAGTTTCTACACCACTTCAATCCCAAGC 5455
Qy 1255 ThrAlaSerGlnLeuSerIleGlnLeuSerSerThrSerThrProThrLeuAlaGlu 1274
Db 5456 ACAGCTTACAGCTGTSCATTCAGCTAGAGAGAGTCTTACTTCTCCTTACCTTGA 5515
Qy 1275 ThrValValAlaSerAlaHisSerLeuAspLysThrSerHisSerSerThrThrGlyLeu 1294
Db 5516 ACCGTGGTAGTACCGACACTCACTAGATTAAGACATCTCAATACAGTACCACTGAGATTG 5575
Qy 1295 AlaPheSerLeuSerAlaProSerSerSerSerSerProGlyAlaGlyValaSerSerTyr 1314
Db 5576 GCTTTCCTCCTCTCTGACCACTTCTCTTCTCTCTGAGGAGAGTGTCTAGTTAT 5635
Qy 1315 IleSerGlnProGlyGlyLeuHisProLeuValIleProSerValIleAlaSerThrPro 1334
Db 5636 ATTTCAGCGCTGGTGGGTGCATCCTTGGTCACTTCATCAGAGTAAGCTTCTACTCCT 5695
Qy 1335 IleLeuProGlnAlaAlaGlyProThrSerThrProLeuLeuProGlnValProSerIle 1354
Db 5696 ATTCCTCCCAAGAGAGAGACCTCACTTCAACCTTATTATCCCAAGTACCTAGTATC 5755
Qy 1355 ProProLeuValGlnProValAlaAsnValProAlaValGlnGlnThrLeuIleHisSer 1374
Db 5756 CCACTCTGTGTAAGCTGTGTCATGTCATGCTGTGTACAGAGACACTAATTCATATGT 5815
Qy 1375 GlnProGlnProAlaLeuLeuProAsnGlnProHisThrHisCysProGlnValaAspSer 1394
Db 5816 CAGCTTCAACCGAGTTGCTTCCCAACCAAGCCCATCTCATTTGTCTGTGAAGTATCT 5875
Qy 1395 AspThrGlnProLysAlaProGlyIleAspAspIleLysThrLeuGlnGlyLysLeuArg 1414
Db 5876 GATACACAAACCAAGCTCTGGAATTGATGACATTAAGACCTAGAGAAAGAGCGCG 5935
Qy 1415 SerLeuPheSerGlnHisSerSerSerSerGlyAlaGlnHisAlaSerValSerLeuGlnThr 1434
Db 5936 TCTCTGTTCAGTAAACAGACTCATCTGAGCTGACATGCTGTGTCTACTGAGAGACC 5995
Qy 1435 SerLeuValIleGlnSerThrValThrProGlyIleProThrThrAlaValAlaProSer 1454
Db 5996 TCACCTAGTCACTAGAGACATGTCACACCAAGCATCCCACTGCTGTGTGACCAAGC 6055
Qy 1455 LysLeuLeuThrSerThrThrSerThrCysLeuProProThrAsnLeuProLeuGlyThr 1474
Db 6056 AAACCTCTGACTTCAACCAAGTACTGCTTCTTACCAACCAATTTACACATAGGAACA 6115
Qy 1475 ValAlaLeuProValThrProValValThrProGlyGlnValSerThrProValaSerThr 1494
Db 6116 GTTGCTTTGGCACTTACACCAAGTGTACACCTGGGCAAGTTCTTACCCAGTCACACT 6175
Qy 1495 ThrThrSerGlyValLysProGlyThrAlaProSerLysProProLeuThrLysAlaPro 1514
Db 6176 ACTACATCAGGAGTGAAGAACTGGAACCTGCTCCTCCAGGCACTCTTACTAAGGCTCCG 6235
Qy 1515 ValLeuProValaGlyThrGlnLeuProAlaGlyThrLeuProSerGlnGlnLeuProPro 1534
Db 6236 GTGCTCCAGTGGGTACTGAATCTTCAGAGTACTTACCCAGCAGAGCTGCCACCT 6295
Qy 1535 PheProGlyProSerLeuThrGlnSerGlnGlnProLeuGlnLysAspAlaGlnLeu 1554
Db 6296 TTTCCAGAGACTTCTTAAACCCAGTCCCAAGCAACCTTGAAGATCTTGAATGCTCAATTG 6355
Qy 1555 ArgArgThrLeuSerProGlu**IleThrValThrSerAlaValGlyProValaSerMet 1574
Db 6356 AGAAGAACACTTAGTCCAGAGATKATCAGATGACTTTCGGTTGGTCTGTGTCTCAAG 6415
Qy 1575 AlaAlaProThrAlaIleThrGlnAlaGlyThrGlnProGlnLysGlyValaSerGlnVal 1594
Db 6416 GCGGCTCYAACAGCATCAACGAAAGAGAAACAACACTCAGAAAGGTGTTTCTCAAGTC 6475
Qy 1595 LysGlnGlyProValaLeuAlaThrSerSerGlyAlaGlyValaPheLysMetGlyArgPhe 1614
Db 6476 AAAAGAGGCTCTGTCTTACGACACTAGTCAAGAGTGTGTGTTTAAAGATGGAGCAATTT 6535
Qy 1615 GlnValSerValAlaAlaAspGlyAlaGlnLysGlnGlyLysAsnLysSerGlnAspAla 1634
Db 6536 CAGGTTCTCTTTCAGAGACAGGCTGCCCAAAAGAGGTAAAAATTAAGTCAAGAAATGCA 6595
Qy 1635 LysSerValHisPheGlnSerSerThrSerGlnSerSerValLeuSerSerSerPro 1654
Db 6596 AAGTGTGTTATTTGAAATCCAGCACTCAGAGTCTCAAGTGTATCAAGTAAAGTACCA 6655
Qy 1655 GlnSerThrLeuValLysProGlnProAsnGlyIleThrIleProGlyIleSerSerAsp 1674
Db 6656 GAGAGTACCTTGTGTGAACCAAGAGCCGAATGGATACCATCCTGTGTATCTTTCAGAT 6715
Qy 1675 ValProGlnSerAlaHisLysThrThrAlaSerGlnAlaLysSerSerThrGlnPro 1694
Db 6716 GTGCCAGAGGTGCCCAAAACTACTGCTCTCAGAGCAAGTCAAGACATCGGAGGCT 6775
Qy 1695 ThrLysValGlyArgPheGlnValThrThrThrAlaAsnLysValGlyArgPheSerVal 1714
Db 6776 ACCAAGTTGACGTTTTCAGTGAACAATCAAGCAAAAGAGTGGTCTTCTCTGTA 6835
Qy 1715 SerLysThrGlnAspLysIleThrAspThrLysLysGlnGlyProValAlaSerProPro 1734
Db 6836 TCAAAAACCTGAGAGACAAAGTCACTGACACAAAGAAAGAGACAGTGGCATCTCCTCT 6895
Qy 1735 PheMetAspLeuGlnGlnAlaValaLeuProAlaValIleProLysLysGlnLysProGlu 1754
Db 6896 TTTATGATTTGGAACAAGCTGTCTTCTGCTGTGTATCCAAAGAAAGAAAGGCTGAA 6955
Qy 1755 LeuSerGlnProSerHisLysAsnGlyProSerSerAspProGluAlaAlaPheLeuSer 1774
Db 6956 CTGTCAAGGCTTCAATCAATGAGGCGCTTCTGACCCGAGGCGCTTTTATAGT 7015
Qy 1775 ArgAspValAspAspGlySerGlySerProHisSerProHisLysSerSerLysSer 1794
Db 7016 AGGATGTGAATGATGATGTCGGTATGCCACATCCGCCATCACTGAGCTCAAGAGC 7075
Qy 1795 LeuProSerGlnAsnLeuSerGlnSerLeuSerAsnSerPheAsnSerSerTyrMetSer 1814
Db 7076 CTTCCTAGCCAGAACTTAAGTCAAAACCTTAGTAATTCATTAACTCTTCTTCAAGAGT 7135
Qy 1815 SerAspAsnGlnSerAspIleGlnAspGlnAspLeuLysLeuLeuArgArgLeuArg 1834
Db 7136 AGGCAATAGAGTCAATATCGAAGATGAAGAACTTAAGTTAAGTGGAGCAACATGCA 7195
Qy 1835 AspLysHisLeuLysGlnIleGlnAspLeuGlnSerArgGlnLysHisGlnIleGlnSer 1854
Db 7196 GATAAACATCTCAAAAGAGATTCAAGACCTGCAGAGTCCAGAAAGCATGAATTTGAATCT 7255
Qy 1855 LeuTyrThrLysLeuGlyLysValaProProAlaValIleIleProProAlaAlaProLeu 1874
Db 7256 TTGTATACCAAACTGGGCAAGGTGCCCTGCTGTATTAATTCCTCCAGTGTCTCCCTT 7315
Qy 1875 SerGlyArgArgArgArgProThrLysSerLysGlySerLysSerSerArgSerSerSer 1894
Db 7316 TCAGGAGAAAGAGCAAGCACTTAATAAAGCAAGCAAAATCTGTAAGCAAGTTC 7375
Qy 1895 LeuGlyAsnLysSerProGlnLeuSerGlyAsnLeuSerGlyGlnSerAlaAlaSerVal 1914
Db 7376 TTGGGAAATTAAGCCCTCAGCTTTCAGGTAACTGTCTGTGTGTGTGTGTGTGTGTGTGT 7435
Qy 1915 LeuHisProGlnIleThrLeuHisIleProProGlyAsnIleProGlnSerGlyGlnAsnGln 1934
Db 7436 TTGCACCCCAAGACAGCTTCAACCTCTGTGGCAATCTCCAGAGTCCGGGAGAAATCAG 7495
Qy 1935 LeuLeuGlnProLeuLysProSerProSerSerAspAsnLeuLysSerAlaPheThrSer 1954
Db 7496 CTGTACAGCCCTTAAAGCATCTCCTCCAGTGAACAATCTATTCACCTTCAACCAAT 7555
Qy 1955 AspGlyAlaIleSerValProSerLeuSerAlaProGlyGlnGlyThrSerSerThrAsn 1974
Db 7556 GATGTGCACTTTCAGTACCAAGCTTCTGTCTCAGAGTCAAGGTAAATTAAGCAACCATC 7615

QY 1975 ThrValGlyAlaThrValAsnSerGlnAlaAla 1985
 Db 7616 ATCGTCCAAAACATATAATGAGATGTTGCC 7648

RESULT 13
 US-10-767-471-449
 ; Sequence 449, Application US/10767471
 ; GENERAL INFORMATION:
 ; APPLICANT: CARGILL, Michele et al.
 ; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
 ; FILE REFERENCE: C1001505
 ; CURRENT APPLICATION NUMBER: US/10/767,471
 ; CURRENT FILING DATE: 2004-01-30
 ; NUMBER OF SEQ ID NOS: 50231
 ; SOFTWARE: PASTESEQ for Windows Version 4.0
 ; SEQ ID NO 449
 ; LENGTH: 9417
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (1)...(9417)
 ; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1-
 US-10-767-471-449)

Alignment Scores:
 Pred. No.: 0 Length: 9417
 Score: 9737.00 Matches: 1963
 Percent Similarity: 88.03% Conservative: 1
 Best Local Similarity: 87.99% Mismatches: 21
 Query Match: 90.06% Indels: 246
 Gaps: 1

US-10-010-720-14 (1-2136) x US-10-767-471-449 (1-9417)

QY 1 MetSerGlyGlyAlaAlaGlnGlnSerSerThrProGlySerLeuPheLeuSerPro 20
 Db 956 ATCTCTGGCGCGCGCGAGAGACAGACAGACACTCCCGTTCCCTGTTCTCGCGC 1015

QY 21 ProAlaProAlaProAlaProAlaGlnGlnSerSerSerSerSerSerValGlyGlnGlnGly 40
 Db 1016 CCGGCTCTCTCCCGCCCAAGATGAGCTCCGATTCCTCCGCGGAGAAATCGGGA 1075

QY 41 AlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 60
 Db 1076 GCGCGCGCGCGCGCGCTGACCGGCGAGACGAGAGTACAGGCGCGCGCGCACACT 1135

QY 61 MetAspLysAspSerArgGlyAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 80
 Db 1136 ATGAGCAAGACAGACCGCTGGGCGCGCGCGCACACTACCACTGAGCAACCGCTTCTTC 1195

QY 81 ArgArgSerValIleCysAspSerAsnAlaThrAlaLeuGlnLeuProGlyLeuProLeu 100
 Db 1196 CGCGGAGCGCTCATCTGNGACTCCCAATGCCACATGCACTGAGCTTCCGCGCTTCTCTT 1255

QY 101 SerLeuProGlnProSerIleProAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 120
 Db 1256 TCCCTGCGCGCGCGCGCGCTGCGCGCGCTGCGCGCGCTGCGCGCTGCGCGCTGCGCGCT 1315

QY 121 ArgGlnGlnThrValThrAlaThrAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 140
 Db 1316 CGGGAAGACGCGTGAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 1375

QY 141 AlaAlaAlaProGlyGlnGlnAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 160
 Db 1376 RCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1435

QY 161 SerLysAspArgProValSerGlnProSerLeuValGlySerLysGlnGlnGlnGlnGln 180
 Db 1436 AGCAAAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1495

QY 181 AlaArgSerGlySerGlyGlySerAlaAlaGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 200

Db 1496 GCGAGAAATGCGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1555

QY 201 GlnAspAspIleGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 220

Db 1556 CAGGATGATATCGAAGAGCTGGAGCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1615

QY 221 LeuLysPheAspIleGlnIleGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 240

Db 1616 CTCAGTTTGAATGCAATTCGAGAGAGCTCTTTAAGCGGTCTACCAAGTCTGAGC 1675

QY 241 ThrGlnThrValGlnAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 260

Db 1676 ACTGAACACCGGTGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1735

QY 261 GlnArgGlnArgPheLysGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 280

Db 1736 GAGAGCGAGAGATTTAAAGAGAGCTGAATGTAAGTCTTCAAGCATCCCAATATT 1795

QY 281 ValArgPheTyrAspSerTrpGlnSerThrValLysGlyLysCysIleValAlaLeuVal 300

Db 1796 GTTAGATTATATATCTCTGGAGATCCAGTAAAGAGAGAGTGTCTTGTGCTG 1855

QY 301 ThrGlnLeuMetThrSerGlyThrLeuLysThrTyrLeuLysArgPheLysValMetLys 320

Db 1856 ACTGAACCTATGACGCTGGAGACCTTAAACGTAATCTGAAAGCTTTAAAGTAGTAG 1915

QY 321 IleLysValLeuArgSerTrpCysArgGlnIleLeuLysGlyLeuGlnPheLeuMetLys 340

Db 1916 ATCAAGTTTCTAAGAGCTGGTGGCGCGCTCAGATCTTAAAGTCTTCAAGTTCTTCAACT 1975

QY 341 ArgThrProProIleIleLysArgPheLysCysAspAsnIlePheIleThrGlyPro 360

Db 1976 CGAATCTCCACTTCACTTACCGCGCATCTTAATGAGACACATCTTATCACCGGCGCT 2035

QY 361 ThrGlySerValLysIleGlyAspLeuGlyLeuAlaThrLeuLysArgAlaSerPheAla 380

Db 2036 ACTGGCTCAGTCAAGATGAGACCTCGCTGCGCAACCTGAAGCGGCGCTTCTTTGCC 2095

QY 381 LysSerValIleGlyThrProGlnPheMetAlaProGlnMetLysGlnGlnGlnGlnGln 400

Db 2096 AAGAGTGTATAGATGACCCGAGAGTTCATGCGCGCGAGTGTATAGAGGAATATGAT 2155

QY 401 GluSerValAspValTyrAlaPheGlyMetCysMetLeuGlnMetAlaThrSerGlyTyr 420

Db 2156 GAATCCGTTGACCTTTATGCTTTGGATGTGATGCTTGAATGCTATCTGAATAT 2215

QY 421 ProLysSerGlnCysGlnAsnAlaAlaGlnIleTyrArgArgValThrSerGlyValLys 440

Db 2216 CTTACTCGGAGTGCAGAAATGCTGCRAGATCTACCGTGGCGTGCACATGCGGTGAG 2275

QY 441 ProAlaSerPheAspLysValAlaIleProGlnValLysGlnIleIleGlnGlnGlnGln 460

Db 2276 CCAGCCAGTTTACAAAGTAGCAATTCGAAAGTAGAAGAAATATTAGAAAGTAGCAT 2335

QY 461 ArgGlnAsnLysAspGlnArgTyrSerIleLysAspLeuLeuAsnIleAlaPheGln 480

Db 2336 CGACAAACCAAGATGAAATATTCATCAAGACCTTTGAACATGCTTCTTCCAA 2395

QY 481 GlnGlnThrGlyValArgValGlnLeuAlaGlnGlnAspGlyGlyLysIleAlaIle 500

Db 2396 GAGGAAACAGAGATAGCGGTAGATTTAGCAGAAAGAGATGAGAGAAAAAATAGCAT 2455

QY 501 LysLeuTrpLeuArgIleGlnAspIleValLysLeuLysGlyLysTyrLysAspAsnGln 520

Db 2456 AAATATATGCTAGATTAAGATATTAAAGAAATTTAAAGGAAATTAACAAAGATATGAA 2515

QY 521 AlaIleGlnPheSerPheAspLeuGlnArgAspValProGlnAspValAlaGlnGlnMet 540

Db 2516 GCTATTAGTTTCTTTGATTTTGAAGAGAGATGCTCCCAAGATGTGCAAGAAATG 2575

QY 541 ValGlnSerGlyTyrValCysGlnGlyAspHisLysThrMetAlaLysAlaIleLysAsp 560

Db 2576 GTAGAGTGGGTATGTCTGTGAAGGTGATCAAGACCATGGCTAAAGCTATCAAGAC 2635
QY 561 ArgValSerLeuIleuValArgLysArgLysGlnArgGlnLeuValArgGlnGlnGln 580
Db 2636 AAGATATCATTAATTAAGAGAAACGAGAGAGCGAGTTGGTACGGAGAGACAAAGAA 2695
QY 581 LysLysLysGlnGlnGlnSerSerLeuLysGlnGlnValGlnGlnSerSerLysGln 600
Db 2696 AAAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2755
QY 601 ThrGlyIleLysGlnLeuProSerAlaSerThrGlyIleProThrAlaSerThrThrSer 620
Db 2756 ACAGAAATCAAGAGCTCCCTCTGTGAGACCGGCACTACTGCTTCTTACCACTTCA 2815
QY 621 AlaSerValSerThrGlnValGlnProGlnGlnProGlnLysAspGlnHisGlnGlnLeu 640
Db 2816 GCTTCAGTTTCTACACAGTAAAGAACTGAGAGAACTGAGGCAATCAACTCAACAACTA 2875
QY 641 GlnTyrGlnGlnProSerLysSerValLeuSerAspGlyThrValAspSerGlyGlnGly 660
Db 2876 CAGTACCGAGACCCAGTATATCTGTGTATCTGATGGACGGTTGACAGTGGTCAAGGA 2935
QY 661 SerSerValPheThrGlnSerArgValSerSerGlnGlnThrValSerTyrGlySerGln 680
Db 2936 TCTCTGTCTTCAVAGATCTCAAGTGAAGACCCAAAGACAGTTCATATGGTTCCCA 2995
QY 681 HisGlnGlnAlaHisSerThrGlyThrValProGlnHisIleProSerThrValGlnAla 700
Db 2996 CATGAACAGGACATTTCTACAGGACAGTCCAGGGCATATACCTTCTACTGTCACAAAGCA 3055
QY 701 GlnSerGlnProHisGlyValTyrProProSerSerValAlaGlnGlnGlnSerGlnGly 720
Db 3056 CAGTCTCAGCCCATGGGATATATCCACCTCAAGTGGCACAGGGGACAGAGCGGT 3115
QY 721 GlnProSerSerSerSerLeuThrGlnValSerSerSerGlnProIleGlnHisProGln 740
Db 3116 CAGCATCTCAAGAGAGCTTAACAGGGGTTTCATCTTCCCAACCCATCAACAACTCTCAG 3175
QY 741 GlnGlnGlnGlnIleGlnGlnThrAlaProProGlnGlnThrValGlnTyrSerLeuSer 760
Db 3176 CAGCAGAGAGGAAATACAGAGACAGGCCCTCTCAACAGACAGTTCATTTCACTTCA 3235
QY 761 GlnThrSerThrSerSerGlnAlaThrThrAlaGlnProValSerGlnProGlnAlaPro 780
Db 3236 CAGACATCAACCTCCAGTGAAGCCACTGACACAGCAGTGAAGTCACTCAAGCTCCA 3295
QY 781 GlnValLeuProGlnValSerAlaGlyLysGln----- 791
Db 3296 CAAAGTCTGCTCAAGATATCAGCTGGAAAACAGCTTCCAGTTTCCAGCCAGTACCAACT 3355
QY 791 ----- 791
Db 3356 ATCCAAAGGGAACCTCAGATCCAGTTGGACACAAACCTCGGTTGTTCAGTCCACTCT 3415
QY 791 ----- 791
Db 3416 GGTGCTCATTTCTTCCAGTGGAGACAGCGCTCCCTACTCTCTTGCTTCTCAGTACCT 3475
QY 791 ----- 791
Db 3476 GTCTTCARATTTCCCATATCACTCTCATGTGTCTACGGGTCAGACAGAGTTTCTCATCC 3535
QY 791 ----- 791
Db 3536 CTTCCCATCAATTTGGAGAGTGGCATTTACTCAGCCTCTGTCTACAGTTGGCTTATCTGCT 3595
QY 791 ----- 791
Db 3596 ACAACAGCTGGCATCCCGGGGTATCAACTGTGGTTCTTAGTACAGTTCCAACTTTCTG 3655
QY 791 ----- 791
Db 3656 CAGCCTGTGACTAGCTGCCAAGTCAAGTTTCAACCCACAGCTCTTACAAACGAGATTTCAG 3715

QY 791 ----- 791
Db 3716 TCATGGGAATACAGCTAACCTTGGACAAAGCTGTGAGTTCCACTTCTTGAGAT 3775
QY 791 ----- 791
Db 3776 GTTCTGACAGGGCTTCCCACTGCACTGCCAACAGTACCCAGAGATTCATATAT 3835
QY 791 ----- 791
Db 3836 GCTCCCTCTTCAACGTGGCTGTGTGATCATTCATTAAGTCTTATCCCTCCATG 3895
QY 791 ----- 791
Db 3896 CCGACAGAGTACTGGCTACCTGGGTACTTTCCACAGTGTGACGCTTATGTGAA 3955
QY 791 ----- 791
Db 3956 TCAATCTTTAGTCTTATGGTGTGTAGAGAGACAGTTCAAGTCCAGCCAGGA 4015
QY 792 ----- 792
Db 4016 GGGAGTTTACACAGACCCCACTACATCTCCAGCAAGCAAGTTTGGAGAGTACTCAG 4075
QY 795 GlyValSerGlnValAlaProAlaGlnProValAlaValAlaGlnProGlnAlaThrGln 814
Db 4076 GAGTCTCTCAGGTTGCTCTCTGAGAGCCAGTTGCAATGACACAGACCAAGCTACCCAG 4135
QY 815 ProThrThrLeuAlaSerSerValAspSerAlaHisSerAspValAlaSerGlyMetSer 834
Db 4136 CCGACCACTTGGCTTCTCTGTGACAGTGCACATTCAGATGTTGGTCAAGTATAGT 4195
QY 835 AspGlyAsnGlnAsnValProSerSerSerGlyArgHisGlnGlnGlnGlnGlnGlnGln 854
Db 4196 GATGGCAATGAGAACGTCCCATCTTCAGTGAAGGACAGAAAGCAAGAAAGAACTCAACAAACGG 4255
QY 855 HisTyrArgLysSerValArgSerArgSerArgHisGlnTyrThrSerArgProLysLeu 874
Db 4256 CATTACCGAAATCTGTAAGAGTCCCTCTCGACATGAAATAAATTCTCAGCCCAAAATTA 4315
QY 875 ArgIleLeuAsnValSerAsnLysGlyAspArgValAlaGlnCysGlnLeuGlnThrHis 894
Db 4316 AGAATTTTGAATGTTCAATAAAGAGACCGAGTGAAGTGAATTAAGTGAAGACTCAT 4375
QY 895 AsnArgLysMetValThrPheLysPheAspLeuAspGlyAspAsnProGlnGlnIleAla 914
Db 4376 AATAGAAAATGGTTACATTCAATTTGACCTTAGTGTGACAAACCCGAGGAGATAGCA 4435
QY 915 ThrIleMetValAsnAsnAspPheIleLeuAlaIleGlnArgGlnSerPheValAspGln 934
Db 4436 ACAATTTATGTAACAATGACTTATTTCTAGCAATGAGAGAGAGTCTTTGTGATCA 4495
QY 935 ValArgGlnIleIleGlnLysAlaAspGlnMetLeuSerGlnAspValSerValGlnPro 954
Db 4496 GTGCGAGAAATTAATGAAAAAGCTGTAAGAAATCTCAGTGAAGATGTCACTGGAACCA 4555
QY 955 GlnGlyAspGlnGlnLysLeuGlnSerLeuGlnGlnLysAspAspTyrGlyPheSerGlySer 974
Db 4556 GAGGGTGAATCAGGAGATTGAGAGTCTACAGGAAAAGATGACATATGCTTTTGAAGTCT 4615
QY 975 GlnLysLeuGlnGlnGlnPheLysGlnProIleProAlaSerSerMetProGlnGlnIle 994
Db 4616 CAGAAATTTGGAAGAGAGATTCACAAACAACCAATTCCTGCTTCCATGCCACGCAATA 4675
QY 995 GlyIleProThrSerSerLeuThrGlnValAlaHisSerAlaGlyArgArgPheIleVal 1014
Db 4676 GGCATTTCTTACCAAGTTCTTAACCTCAAGTTGTCTTACGCGGGAAGCGGTTTATAGTG 4735
QY 1015 SerProValProGlnSerArgLeuArgGlnSerLysValPheProSerGlnIleThrAsp 1034
Db 4736 AGTCTGTGACGAAAAGCGAATTTACGAGATCAAAAGTTTCCCGCAGTGAATTAACAGAT 4795

QY	1035	ThrValAlaAlaIaSerThrAlaGlnSerProGlyMetAsnLeuSerHisSerAlaSerSer	1054
Db	4796	ACAGTGTGCTCTTACAGCTCAGAGCCCTGGAATGAACCTTGCTCACCTTGCAATCAACC	4855
QY	1055	LeuSerLeuGlnGlnAlaPheSerGluLeuLeuArgAlaGlnMetThrGluGlyProAsn	1074
Db	4856	CTTAGCTCAACAACAGCCTTTTCTGAACTTAGACTGCCCCAATACAGAAAGAACCAAN	4915
QY	1075	ThrAlaProProAsnPheSerHisThrGlyProThrPheProValValProPheLeu	1094
Db	4916	ACAGAACCTCCAAACTTAGTCATCATCAGAGCAACCAATTCAGATGATCTCTTCTTA	4975
QY	1095	SerSerIleAlaGlyValProThrThrAlaAlaAlaThrAlaProValProAlaThrSer	1114
Db	4976	AGTACATATGCTGGAGTCCCAACACACAGACAGACAGACACAGCTCCCTGCACAACGC	5035
QY	1115	SerProProAsnAspIleSerThrSerValIleGlnSerGluValThrValProThrGlu	1134
Db	5036	AGCCTCCTTAATGACATTTCCACATCAGTAATTCGTGAGTTACAGTGCCTGAA	5095
QY	1135	GluGlyIleAlaGlyValAlaAlaThrSerThrGlyValValThrSerGlyLeuProIle	1154
Db	5096	GAGGGGAATGCTGGAGTGCACACAGACAGAGGTGGTAATCTTACAGGTGCTCCCCATA	5155
QY	1155	ProProValSerGlnSerProValLeuSerSerValValSerSerIleThrIleProAla	1174
Db	5156	CCACCTGTCTGGAATCACACAGTAACTTTCACGCGTAGTTACAGATACAAATACCTGCA	5215
QY	1175	ValValSerIleSerThrThrSerProSerLeuGlnValProThrSerThrSerGluIle	1194
Db	5216	GTTGCTCAAAATCTACTACATCCCCGTCACCTTCAGAGTCCCCCAATCCACATCTGAGATC	5275
QY	1195	ValValSerSerThrAlaLeuLeuLeuProSerValThrValSerAlaThrSerAlaSerAla	1214
Db	5276	GTTGTTTCTAGTACACAGCACTGATCTTCAGTAAAGTTTCAGACAACTTCAGCCTTGCA	5335
QY	1215	GlyGlySerThrThrAlaThrProGlyProLysProProAlaValValSerGlnAlaAla	1234
Db	5336	GGGGCACTACTGCTCAACCCACAGTCCCTTAAGCCTCAGCTGTAAGATCCACAGGCACACA	5395
QY	1235	GlySerThrThrValGlyAlaAlaThrLeuThrSerValSerThrThrThrSerPheProSer	1254
Db	5396	GGCAGCACTACTGTGGGAGGCCAATTAAACATCAGTTTTCACCACTTCATTTCCCAAGC	5455
QY	1255	ThrAlaSerGlnLeuSerIleGlnLeuSerSerThrThrSerThrProThrLeuAlaGlu	1274
Db	5456	ACAGCTTCACAGCTGTSCATTCAGCTTAGCAGCACTACTTCACTCACTTACTTACTGATAA	5515
QY	1275	ThrValValValSerAlaHisSerLeuAspLysThrSerHisSerSerThrThrGlyLeu	1294
Db	5516	ACCGGTGATGATGAGGCACACTCACTCAATATAGACATCTCATAGCACTACCAACTGGATTGG	5575
QY	1295	AlaPheSerLeuSerAlaProSerSerSerSerSerProGlyAlaGlyValSerSerLys	1314
Db	5576	GCTTTCTCTCTCTGCAACCAATCTTCTCTCTCTCTCTGAGAGGAGGTGTCTAGTAT	5635
QY	1315	IleSerGlnProGlyGlyLeuHisProLeuValIleProSerValIleAlaSerThrPro	1334
Db	5636	ATTTCAGCTGCGGTGGGCTGTGATCCTTTGGTCATTCATCAGTATAGCTTCACTACTCT	5695
QY	1335	IleLeuProGlnAlaAlaGlyProThrSerThrProLeuLeuProGlnValProSerIle	1354
Db	5696	ATTCTTCCCCAAGCAGCAGCACTCACTCTCAACCTTATTATACCCCAAGTACTGTAGTATC	5755
QY	1355	ProProLeuValGlnProValAlaAsnValProAlaValGlnGlnThrLeuIleHisSer	1374
Db	5756	CCACCTTGTGATCACCTGTTGGCAATATGCTGTGTACAGCAACAACATAATCATAGT	5815
QY	1375	GlnProGlnProAlaLeuLeuProAsnGlnProHisThrHisLysProGlyValAlaSpSer	1394
Db	5816	CAGCCTCAACAGCTTTGCTTCCCAACACAGCCCCCATCTCATTTGCTCTGAAGTACATTTCT	5875
QY	1395	AspThrGlnProLysAlaProGlyIleLeuAspIleLysThrLeuGlnGluLysLeuArg	1414

Db	5876	GATACACACC	AAAGCTCCTGGAATTGATGACATTAAGACTTAGAAGAAAGCTCGG	5935
QY	1415	SeIUeuPheSerGIuHISerSeSerGIyAlaGlnHISaIaSerValSerIeuGIuThr		1434
Db	5936	TCTGTGTTCAATGAACA	CAGCTCATCTGGAGCTCAGCATCCTCTGTCTCATCTGAGAAC	5995
QY	1435	SeIUeuValIIeGIuSerThrValThThProGIyILeProThrThraIaValaIaProSer		1454
Db	5996	TCACCTAGTCATAGAGACAC	ACTGTACACAGGCTCCCACTACTGCTGTGGACCAAGC	6055
QY	1455	LySIeuThrThSerThrThrThSerThrCysIeuProProThraISnIeuProIeuGIyThr		1474
Db	6056	AAACTCTGACTTCTACCA	CAAGTACTTGCTTTACACAAACCAATTTCACATGGAACA	6115
QY	1475	ValaIaIeuProValaThThProValaIaThThProGIyGlnValSerThThProValaSerThr		1494
Db	6116	GTTCTCTTGGCAGTTAC	CAACAGTGTGTACACCTGGGGAAGTTTCTACCCAGTACGACACT	6175
QY	1495	ThThSerGIyValIySProGIyThThaIaProSerIySProProIeuThThIySaIaPro		1514
Db	6176	ACTACATCAGAGAGGAAC	CTGGAACCTCCCTCCAAAGCCACTCTACCTTAAGAGCTCCG	6235
QY	1515	ValIeuProValaGIyThThGIuIeuProValaGIyThThIeuProSerGIuGlnIeuProPro		1534
Db	6236	GTCGTCCAGAGGGGTACTGA	ACTTCACAGAGATCTTACCCACGAGCAGCTGCACCT	6295
QY	1535	PheProGIyProSerIeuThrGlnSerGIyGlnProIeuGIuSaIeuAspIaGlnIeu		1554
Db	6296	TTTCCAGGACCTTCTTA	CCCACTCCACAGAACCTCTAAGAGATCTTGAATGCTCAATTG	6355
QY	1555	ArgArgThrIeuSerProGIu***IeThrValThSerAlaValaGIyProValaSerMet		1574
Db	6356	AGAAGAACA	CTTACAGAGATATATACAGTCACTTCGCGGTGGTCTCTGTTCATG	6415
QY	1575	AlaIaIaProThraIaIeThThGIuIaGIyThThGlnProGlnIySaIeuSerGIuVal		1594
Db	6416	GCGGCTCYAAAGAGAACT	ACACAGGAAGCAACAGCCTCAGAAGGGGTGTTTCAAGTC	6475
QY	1595	LySGlnGIyProValaIeuAlaThSerSeSerGIyAlaGIyValaPhelysMetGIyArgPhe		1614
Db	6476	AAAGAAGGCCCTGCTTA	GCACACTAGTTCAGAGAGCTGTGTTTTAAGATGGAGCATTT	6535
QY	1615	GlnValaSerValaIaIaIaSPeGIyAlaGlnIySaIeuGIyIySaIeuIySaIeuIySaIa		1634
Db	6536	CAGGTTTCTGTTCAGACA	CGGTGCCCAAAAGAGGTAAATATAGTCAAGAAGATGCA	6595
QY	1635	LySerValaHISpHeGIuSerSeThThSerGIuSerValaIeuSerSeSerPro		1654
Db	6596	AAAGTCTGTCTATTGAA	TCCAGCACTCAGAGCTCCAGTGCTATCAAGTAGTAGTCA	6655
QY	1655	GlnSerThrIeuValaIySProGIuProAsnGIyILeThrIleProGIyIleSerSeAsp		1674
Db	6656	GAGAGTACTTGTGGAA	CAACAGACCGAATGGCATTAACATCCTGGTATCTCTTCAGAT	6715
QY	1675	ValProGIuSerAlaHISlyThThThraIaSerGIuAlaIySaIeuAspThThGIyGlnPro		1694
Db	6716	GTGCCAGAGAGTGC	CCCAAACTAGCTCCAGAGGCAAAAGTGCAGACACTGGGCACACT	6775
QY	1695	ThThIySaIaGIyArgPheGlnValThThThThraIaSnIySaIaGIyArgPheSerVal		1714
Db	6776	ACCAAGGTGGACGTTT	CAGGTCAACAATCACAAACAAAGGGGTGCTTCTTCGTA	6835
QY	1715	SeIyThThGIuSaIySILeThAspThThIySaIySGlnGIyProValaIaSerProPro		1734
Db	6836	TCAAAACTGAGGAC	AGATATCTAGACACAAAGAAAGAACCAAGTGGCATCTCTCCT	6895
QY	1735	PheMetAspIeuGIuGlnAlaValaIeuProAlaValIleProIySlySGlnIySProGIu		1754
Db	6896	TTTATGATTTGGAACA	AGCTGTTCTTCTGCTGTGATACCAAGAAAGARAAGCTCGAA	6955
QY	1755	IeuSerGIuProSerHISleuAsnGIyProSerSeAspProGIuAlaIaPheIeuSer		1774

Dd		6956 CTGTACAGACCCTTCAATCATTAATAAGGCGCGTTCTTGACCCGGAGGCCGCTTTTAACT	7015
Oy	1775	ArgAspPvalAspAspGlySerGlySerProHisSerProHisGlnLeuSerSerIysSer	1794
Dd	7016	AGGAGATGGAGTAGATGGTTCCGGTGTGCACACTGTGCCCACTAGACTGAGCTCAAAGAGC	7075
Oy	1795	LeuProSerGlnAanLeuSerGlnSerIleuSerAsnSerPheAsnSerSerTyMetSer	1814
Dd	7076	CTTCCTAGCCGAATCTTAAGTAAAGGCTTGTAGTAATTATTAACTCTCTTACATGAGT	7135
Oy	1815	SerAspAsnGluSerAspIleGlnAspGlnuAspLeuLysLeuGluLeuArgLeuArg	1834
Dd	7136	AGCGACATAGAGTGCATATCGAGAATAAGACTTAAAGCTTAGAGCTCGACGACTACGA	7195
Oy	1835	AapLysHisLeuLysGluIleGlnAspLeuGlnSerArgGlnLysGluIleGluSer	1854
Dd	7196	GATAAACACTCTCAAGAAGATTACGAGACTTCGACAGAGTCGCCAAGACATGAATGAATCT	7255
Oy	1855	LeuTrpThrLysLeuGlyLysValProProAlaValIleIleProProAlaAlaProLeu	1874
Dd	7256	TGTATACCAACTGGGCAAGGTGCCCCCTGCTGTATTATTATTTCCCACGCTGCCCTT	7315
Oy	1875	SerGlyArgArgArgProThrLysSerLysGlySerLysSerSerArgSerSerSer	1894
Dd	7316	TCAGGAGAAACAACGACGACCCACTTAAGACAAAGCAGCAAAATGATGACGACATTC	7375
Oy	1895	LeuGlyAsnLysSerProGlnLeuSerGlyAsnLeuSerGlyGlnSerAlaAlaSerVal	1914
Dd	7376	TTGGGGATTAATAAGCCCCAGCTTCAAGTAACTGTGTGGTCAGAGTGCAGCTTCAGTC	7435
Oy	1915	LeuHisProGlnGlnThreuHisProProGlyAsnLleProGluSerGlyGlnAsnGln	1934
Dd	7436	TTGACACCCCAAGCAACCCCTCACCCCTCTGGCAACATCCACAGAGTCGGGCAAACTAG	7495
Oy	1935	LeuLeuGlnProLeuLysProSerProSerSerAspAsnLeuTyrsenAlaPheThrSer	1954
Dd	7496	CTGTTACGCCCCCTTAAGCAATCTCCCTCCAGTGAACAACCTTAATCAGCTTACCAAGT	7555
Oy	1955	AspGlyAlaIleSerValProSerLeuSerAlaProGlyGlnGlyThrSerThrAsn	1974
Dd	7556	GATGGTGCCATTTCAGTACCAAGCTTCTCTCCAGSTCAAGGTAAATAAGCAACATC	7615
Oy	1975	ThrValGlyAlaThrValAsnSerGlnAlaAla	1985
Dd	7616	ATCGTCCAAAAACAATAAATGAGAGATGTTGCC	7648
RESULT 14			
US-09-979-167-63			
; Sequence 63, Application US/09979167			
; GENERAL INFORMATION:			
; APPLICANT: PLOWMAN, GREGORY D.			
; APPLICANT: MARTINEZ, RICARDO			
; APPLICANT: MYHRE, DAVID			
; APPLICANT: SUBERSANAM, SUCHA			
; TITLE OF INVENTION: PROTEIN KINASES			
; FILE REFERENCE: 038602/1273			
; CURRENT APPLICATION NUMBER: US/09/979,167			
; CURRENT FILING DATE: 2001-11-20			
; NUMBER OF SEQ ID NOS: 269			
; SOFTWARE: PatentIn Ver. 2.1			
; SEQ ID NO 63			
; LENGTH: 7328			
; TYPE: DNA			
; ORGANISM: Homo sapiens			
; FEATURE:			
; NAME/KEY: modified_base			
; LOCATION: (2083)..(2084)			
; OTHER INFORMATION: a, t, c, g, other or unknown			
US-09-979-167-63			
Alignment Scores:			
Pred. No.:	0	Length:	7328
Score:	9729.50	Matches:	1939

[illegible]

Db 1002 ATCAAGTTCATAAGAACTGTCGCCGTCAGATTCCTTAAGTCTTCAGTTCCTTCATCT 1061
QY ArgThrProPro11e11eh1sargaspLeu1yCyaspasn11ePhel1eThrglyPro 360
Db 1062 CGAATCCACTTCATTCATCCCGCATCTTAAATGACAAACATCTTATACCCGGCCCT 1121
QY ThrG1SerVal1y1s1leg1yaspLeu1yLeu1aThrLeu1yAsrg1aSerPhe1a 380
Db 1122 ACTGGCTCAGTCAAGATGAGACCTGCTGTGGCAACCTTGAGCGGCTCTTTGGC 1181
QY LysSerVal11leg1yThrProGluPheMet1aProGluMet1yTyrglu1yTyasp 400
Db 1182 AAGAGTGTAGTATGATCCAGAGTTCATGCCCCCTGATGATGAGGAGAAATATGAT 1241
QY GluSerVal1aspVal1yTyra1aPheGlyMet1yMet1eLeu1uMet1aThrSerGluTy 420
Db 1242 GAATCCGTTGACGTTTATGCTTTGGATGTGCAATGCTTGAATGCTCATCTGAAATAT 1301
QY ProTySerGluCyagl1asn1a1a1aGlu1eTyraArgVal1ThrSerGlyVal1ys 440
Db 1302 CTTTACTCGSAGTCCAAATGCTGGCAGATCTACCGTGGCTGACCACTGGGGTGAG 1361
QY ProAl1SerPheasp1yVal1a11eProGluVal1yGlu1e1leg1yCy1s1e 460
Db 1362 CCAGCCAGTTCACAAAGTAGCAATTCCTGAAGTGAGGAAATATTTGAAGATGCAT 1421
QY ArgGluAsn1yaspGluArgTySer11eLysAspLeu1eAsn1s1a1aPhePheGlu 480
Db 1422 CGACAAACAAAGATGAAATATTCATCAAGACCTTTTGAAACCAATGCTTCCTCA 1481
QY GluGluThrg1yVal1ArgVal1GluLeu1a1Glu1uaspaspGluTy1s1e1a11e 500
Db 1482 GAGGAAACAGAGATCCGGTNGAATTTAGCAAGAAAGATGATGAGAAAAAATAGCCATA 1541
QY LysLeu1eTypleu1yrg11eGluasp11eLysLysLeu1yG1yTyTy1yaspasnGlu 520
Db 1542 AAATTTAGCTGATGATTTGAAGATATTAAGAAATTTAAAGGAAATATCAAGATTAATGA 1601
QY Ala11eGluPheSerPheaspLeuGluArgaspVal1ProGluaspVal1a1aGlu1uMet 540
Db 1602 GCAATGAGTTCGTTGTTGTTTGAAGAGATGTCCTCAAGAAATGTTGACAAAGAAATG 1661
QY ValG1uSerG1yTyra1yCyagl1yAsp11eLysThrMet1a1y1s1a11eLysasp 560
Db 1662 GTAGAGTTCGGATATGCTGTGAAGTGATCAACAAGCCATGGCTAAAGCTATCAAGAAC 1721
QY ArgVal1SerLeu11eLysArg1y1sargGluGluArgGluLeuVal1ArgGluGluGlu 580
Db 1722 AGAGTATCATTAATTTAAGAGAAACGAGACAGCGGATGTTGTTACGGGAGCAAGAA 1781
QY LysLysLysGluGluGluGluSerSerLeu1yGluGluVal1GluGluSerSer1a1aSerGlu 600
Db 1782 AACAAAAGAGAGAGAGAGAGAGCTCAACAGAGATGAGAAACATCACTGCTCCAG 1841
QY ThrG1y11eLysGluLeuPheSer1a1aSerThrg1y11eProThra1aSerThrThSer 620
Db 1842 ACAGGAATCAAGACGCTCCCTTCGCTAGCAACCGCATACCTACTGTTTACACACTCA 1901
QY Al1SerVal1SerThrgluVal1GluProGluGluProGlu1aaspGlu1s1eGlu1uMet 640
Db 1902 GCTTCAGTTCCTACACAGATGAAACCTGAAACCTGAGGACAGTCAACATCAACAATA 1961
QY GluTyrgluGluProSer11eSerVal1LeuSeraspGlyThrVal1aaspserGlyGlu 660
Db 1962 CAGACACAGAACCCAGATATATCTGTTATCTGATGGAGCGGTTGACATGGTCAGGGA 2021
QY SerSerVal1PheThrgluSerArgVal1SerSerGlu1uThrVal1aSerTyrgly1---Ser 679
Db 2022 TCTTCTGTCTTCACAGATCTCGAGTGAAGCAACCAAGTTCATATGGGTTCCCA 2081
QY Glu1h1eGluGlu1a1a1h1sSerThrglyThrVal1ProGly1h1s11eProSerThrVal1Glu 699
Db 2082 ANNCATGAACAGGACATCTTCAAGGACAGTCCCAAGGACATATACCTTCACTGTCCAA 2141

QY AlaGluSerGluPro1h1sGlyVal1TyProProSerSerVal1a1aGluGluGluSerGlu 719
Db 2142 GCACAGTCTCAGCCATGGGATATATCCACCTCAAGTGTG----- 2183
QY GlyGluProSerSerSerSerLeuThrglyVal1SerSerSerGluPro1leg1u1h1sPro 719
Db 2183 ----- 2183
QY GluGluGluGluGlu11eGluGluThra1aProProGluGluThra1aGluTySerLeu 719
Db 2184 -----CAGCAGGAAATACAGCAGACAGCCCTCCCAACAGACAGTGCATTCATCTT 2237
QY SerGluThrSerThrSerSerGlu1a1aThrThra1aGluProVal1SerGluProGlu1a 719
Db 2238 TCACACATCAACCTTCAGTGAAGGCACTACGACACAGCAAGTGAAGTCAAGCTCAAGCT 2297
QY ProGluVal1LeuProGluVal1Ser1aGlyLysGluSerThrgluGluVal1aSerGluVal 719
Db 2298 CCAAGATCTTGCCTCAAGTATCAGCTGGAAAACAGATCTCAGGAGATCTCAGGTT 2357
QY AlaProAl1GluProVal1a1aVal1a1aGluProGlu1a1aThrgluProThrThra1a 819
Db 2358 GCTCCGACAGAGCAGATTCGAGTGAAGCAGCCCAAGCTAACCAAGCCGACCTTGGCT 2417
QY SerSerVal1aSerSerAl1h1sSeraspVal1a1aSerGlyMet1yMet1yAspGluAsn 839
Db 2418 TCTCTGTGACAGTGCATTCAGATGTTGCTTCAGGATGATGATGATGATGATGATGATG 2477
QY Val1ProSerSerSerGlyArg1h1sGluGlyArgThrglyThrglyThrglyThrglySer 859
Db 2478 GTCCCATCTTCAGTGAAGGATGAGGAAAGATACAAAACGCGATTCAGGAAATATC 2537
QY Val1ArgSerThrgSerThrg1h1sGlyTy1sThrSerArgProLysLeuArg11eLeuAsnVal 879
Db 2538 GTAAAGATGCTGCTCTGACATGAAAACTTCACGCCCAAAATTAAGAAATTTGAATGTT 2597
QY SerAsnLysGlyAspArgVal1a1GluCyGluLeuGluThrg1h1sAsnArgLysMetVal 899
Db 2598 TCMAATTAAGAGAACCGAGTATGATGATGATGATGATGATGATGATGATGATGATGATG 2657
QY ThrPheLysPheaspLeuaspGlu1yaspasnProGluGlu1e1aThr11eMetValasn 919
Db 2658 ACATTCAAATTTACCTAGATGCTGACCAACCCGAGGAGATGACCAATTAATGATGATG 2717
QY AsnAspPhe11eLeu1a11eGluArgGluSerPheVal1aspGluVal1a1aGlu1u1e1e 939
Db 2718 AATGACTTATTTCTACAAATGAGAGAGAGATGCTGTTGTGATCAAGTGCAGAAATTAAT 2777
QY GluTyra1aaspGluMetLeuSerGluaspVal1SerVal1GluProGluGlu1yaspGluGly 959
Db 2778 GAAAAAGCTGATATAATTCCTCACTGAGATGTCAAGTGTGAACACAGAGGCTGATCAGGGA 2837
QY LeuGluSerLeuGluGlyLysaspAspTyrglyPheSerGlySerGluTyLeuGluGly 979
Db 2838 TTGAGAGTGTACAAAGAAAGATGATGATGCTTTTCAGGTTCTCAGAAATTTGAAGGA 2897
QY GluPheLysGluPro1eProAl1aSerSerMet1yProGluGlu11eGly11eProThrSer 999
Db 2898 GAGTTCAAACAAACAAATTCCTGCTGCTTCATGCTCAACAGCAAAATGAGCATTCACAGT 2957
QY SerLeuThrgluVal1h1sSerAl1aGlyArgArgPhe11eVal1SerProVal1ProGlu 1019
Db 2958 TCTTAACTCAATGTTGATCTCTGGGGAAGCGGTTTATGATGATCCTGTGCGCAGAA 3017
QY SerArgLeuArgGluSer1yVal1PheProSerGlu11eThraaspThra1a1a1aSer 1039
Db 3018 AGCCGATTAAGAAATCAAAAGTTTCCCAAGGAAATACAGATACAGTGTGCTGCTCT 3077
QY ThrAl1GluSerProGluMetLeuLeuSerH1sSerAl1aSerSerLeuSerLeuGluGlu 1059
Db 3078 ACAGCTCAGAGCCCTGGAAAGAACTGTCTCACTGTGATCATCTTATGTACAAACAG 3137

QY 1060 AlaPheSerGluLeuAlaGagAlaGlnMetThrGluGlyProAsnThrAlaProProAsn 1079
DB 3138 GCCCTTTCCTTAACCTTACAGCTGCCCCAATATGACAGAAAGACCCAAATATACAGACCTCCAAAC 3197
QY 1080 PheSerHisThrGlyProThrPheProValValProPheLeuSerSerIleAlaGly 1099
DB 3198 TTATGATATACAGAACCAACATTTCCAGTAGTACTCTTCTTAACTATACACTTGGCTGA 3257
QY 1100 ValProThrThrAlaAlaAlaThrAlaProValProAlaThrSerSerProProAsn 1119
DB 3258 GTCCCAACCAAGAGAGACCAAGCAAGCAAGCTCCCTGCAACAAAGACCCCTCTAATGAC 3317
QY 1120 IleSerThrSerValIleGlnSerGluValThrValProThrGluGluGlyIleAlaGly 1139
DB 3318 ATTTCACATCAGTAATTCAGTCTGAGGTACAGTGCCCACTGAAGAGGGGATTTGCTGA 3377
QY 1140 ValAlaThrSerThrGlyValValThrSerGlyGlyLeuProIleProProValSerGlu 1159
DB 3378 GTTGCCACCAAGACAGAGGTGTGTACTTCAAGTGTCTCCCAATACCACTGTGTGAA 3437
QY 1160 SerProValLeuSerSerValValSerSerIleThrIleProAlaValValSerIleSer 1179
DB 3438 TCACCAAGTACTTCCAGCGTAGTTTCAAGTATCAACATACCTGCAAGTGTCTCAATATCT 3497
QY 1180 ThrThrSerProSerLeuGlnValProThrSerThrSerGluIleValValSerSerThr 1199
DB 3498 ACTACATCCCCGTCACCTTCAAGTCCCAACATCCACATCTGAGATCGTTGTTCTAGTACA 3557
QY 1200 AlaLeuThrProSerValThrValSerAlaThrSerIleSerAlaGlyValGlySerThrAla 1219
DB 3558 GCACGTATCTTCAGTAACAGTTTCAGCAACTTCAGCTCTGAGGGGAGACTGCT 3617
QY 1220 ThrProGlyProLysProProAlaValValSerGlnAlaIleAlaGlySerThrThrVal 1239
DB 3618 ACCCAAGTCTTACAGCTCCAGCTAGTATCTCAGAGGACAGACAGCACTACTGTC 3677
QY 1240 GlyAlaThrLeuThrSerValSerThrThrThrSerPheProSerThrAlaSerGlnLeu 1259
DB 3678 GAGGCCACATTAATCATGATTTCTACACCACTTCACTCCCAAGCACAGCTTCAAGCTG 3737
QY 1260 SerIleGlnLeuSerSerSerThrSerThrProThrLeuAlaGlnThrValValValSer 1279
DB 3738 TCCATTTAGCTTACAGACAGTACTTCACTCTCACTTAACTGAACCGGTGATGAGTAC 3797
QY 1280 AlaHisSerLeuAspLysThrSerHisSerSerThrThrGlyLeuAlaPheSerLeuSer 1299
DB 3798 GCACACCTCACTAGATTAAGACATCTCATAGCAAGTACAACTGGATTTGCTTCCCTCTC 3857
QY 1300 AlaProSerSerSerSerSerProGlyAlaGlyValSerSerTyrIleSerGlnProGly 1319
DB 3858 GCACCAATCTTCTCTCTCTCTCTGAGAGAGTGTCTAGTTATATTCTCAGCCGTGT 3917
QY 1320 GlyLeuHisProLeuValIleProSerValIleAlaSerThrProIleLeuProGlnAla 1339
DB 3918 GGGCTGATCTTGGTCATTCATCATGATAGCTTCACTCTATCTTCTCCCAAGCA 3977
QY 1340 AlaGlyProThrSerThrProLeuLeuProGlnValProSerIleProProLeuValGln 1359
DB 3978 GCAGGACCTACTTCACTCTTATTTACCCCAAGTACTGTATCCCACTTGGTACAG 4037
QY 1360 ProValAlaAsnValProAlaValGlnGlnThrLeuIleHisSerGlnProGlnProAla 1379
DB 4038 CCTGTGGCCATGTGCTGCTGTACAGCACTAATTCATAGTACGCTCAACACGCT 4097
QY 1380 LeuLeuProAsnGlnProHisThrHisCysProGluValAlaAspSerAspThrGlnProLys 1399
DB 4098 TTACTTCCCAACAGCCCACTCATTTCTGGAATAGATTTCTGATACAAACCCAAA 4157
QY 1400 AlaProGlyIleAspAspIleLysThrLeuGluGluLysLeuArgSerLeuPheSerGlu 1419
DB 4158 GCTCTCGGATTTATGACATTAAGACCTCTGAAGAAAGAAAGCTGGGCTCTCTCAAGTGA 4217
QY 1420 HisSerSerSerGlyAlaGlnHisAlaSerValSerLeuGluThrSerLeuValIleGln 1439

DB 4218 CACAGCTCATCTGGAGCTCAGCATGCTGTCTCACTGGAGACCTCATGATCATGAG 4277
QY 1440 SerThrValThrProGlyIleProThrThrAlaValAlaProSerLysLeuLeuThrSer 1459
DB 4278 AGCAGTGTACACAGGACATCCCACTACTGCTGTGGACCAAGCAAACTCTGACTTCT 4337
QY 1460 ThrThrSerThrCysLeuProProThrAsnLeuProLeuGlyThrValAlaLeuProVal 1479
DB 4338 ACCACAAGTACTTGTCTTACCAACCAATTTTCCATAGAACAGTGTGTTGGCCAGTT 4397
QY 1480 ThrProValValThrProGlyGlnValSerThrProValSerThrThrThrSerGlyVal 1499
DB 4398 ACACCAAGTGTCACTCTGGGCAAGTTTTCACCCCACTGACACTCAATCATCAGAGAGT 4457
QY 1500 LysProGlyThrAlaProSerLysProProLeuThrLysAlaProValLeuProValGly 1519
DB 4458 AAACCTGGAACCTCTCTCCCAAGCACTTAACTGAAGGCTCCGGTGTGCCAGTGGGT 4517
QY 1520 ThrGluLeuProAlaGlyThrLeuProSerGluGlnLeuProPheProGlyProSer 1539
DB 4518 ACTGAATCTCCAGCAAGTACTTACCCAGAGACAGCTGCCACTTTTCCAGACACTTCT 4577
QY 1540 LeuThrGlnSerGlnProLeuGluAspLeuAspAlaGlnLeuArgArgThrLeuSer 1559
DB 4578 CTAAACCAAGTCCCAAGCACTCTAGAGAGATCTTGATGCTCAATTGAGAAAGAACCTTAGT 4637
QY 1560 ProGlu**IleThrAlaThrSerAlaValAlaGlyProValSerMetAlaAlaProThrAla 1579
DB 4638 CCAAGATGATTCACAGTGAATCTTGGGTGTGTGTGTCCATGGGGGTCCCAACAGCA 4697
QY 1580 IleThrGluAlaGlyThrGlnProGlnLysGlyValSerGlnValLysGluGlyProVal 1599
DB 4698 ATCACAAGAAAGCAAGCAAGCTCAGAAAGGATTTTCAATCAATAAGAAAGCCCTGTC 4757
QY 1600 LeuAlaThrSerSerGlyAlaGlyValPheLysMetGlyArgPheGlnValSerValAla 1619
DB 4758 CTAGCAACTTGAATTCAGAGAGTGTGTTTAAATATGAGCAGATTTCAAGTTCTGTTGCA 4817
QY 1620 AlaAspGlyAlaGlnLysGluGlyLysAsnLysSerGluAspAlaLysSerValHisPhe 1639
DB 4818 GCAGAGGTGCCCCAAGAAAGGTTAAATTAATTCGAAGATCAAAAGTCTGTTCATTCT 4877
QY 1640 GluSerSerThrSerGluSerSerValLeuSerSerSerSerProGluSerThrLeuVal 1659
DB 4878 GAATCCAGACCTCAGAGTCTCAGTGTATCAAGTAGTAGTCCAGAGATACTTGCTGTG 4937
QY 1660 LysProGluProAsnGlyIleThrIleProGlyIleSerSerAspValProGluSerAla 1679
DB 4938 AAACCAAGAGCCGAATGCAATACCATCCCTGATATCTTCAAGATGTGCCAGAGAGTGC 4997
QY 1680 HisLysThrThrAlaSerGluAlaLysSerAspThrGlyGlnProThrLysValGlyArg 1699
DB 4998 CACAAATCTATGCTCTCAAGGCAAGAGTCAAGACCTGAGGACCTTCAAGGTTGACGT 5057
QY 1700 PheGlnValThrThrThrAlaLeuLysValGlyArgPheSerValSerLysThrGluAsp 1719
DB 5058 TTTCAGGTGACATCTACAGAAACAAAGTGGTCTTCTCTGTATCAAAAATCGAGAC 5117
QY 1720 LysIleThrAspThrLysLysGluGlyProValAlaSerProProPheLeuAspLeuGlu 1739
DB 5118 AAGATCACTGACACAAAGAAAGAGACAGTGGCATCTCTCTTTATGTGATTGGAA 5177
QY 1740 GlnAlaValLeuProAlaValIleProLysLysGluLysProGluLeuSerGluProSer 1759
DB 5178 CAAGCTGTCTTCTCTGTGATACCAAGAAAGAGGCTTAAGTGTAGAGCTTCA 5237
QY 1760 HisLeuAsnGlyProSerSerAspProGluAlaAlaPheLeuSerArgAspValAspAsp 1779
DB 5238 CATCTAAATGGGCGCTCTTCTGACCCGAGAGCGCTTTTAAAGTAGGATGTGATGAT 5297
QY 1780 GlySerGlySerProHisSerProHisGlnLeuSerSerLysSerLeuProSerGlnAsn 1799

Db 5298 GGTTCCGAGTAGTCACACTCGCCCATCAGCTGACGCTCAAGAGCCCTTCTAGCCAGAAAT 5357
 QY 1800 LeuSerGlnSerLeuSerAsnSerPheAsnSerSerTyneSerSerAspAsnGluSer 1819
 Db 5358 CTAGATGAAGCCCTAGTAATTCATTATTAACCTCTTACATGAGTACGCAATGAGCTCA 5417
 QY 1820 AspIleGlnAspGluAspLeuLeuGlnLeuLeuArgArgLeuAspAspLysHisLeuLys 1839
 Db 5418 GATTCGAAAGTAGAGACTTAAAGTTAGAGTGGGAGGACGACGACGATTAACATCTCAA 5477
 QY 1840 GlnIleGlnAspLeuGlnSerArgGlnLysHisGlnIleGlnSerLeuTyThrLysLeu 1859
 Db 5478 GAGATTGAGAGCCGCGAGAGCGCCAGAGCATGAATGAACTTTGATACCAACTG 5537
 QY 1860 GlyLysValProProAlaValIleIleProProAlaAlaProLeuSerGlyValArgArg 1879
 Db 5538 GGCAGAGGTGCCCCGTGTTATTTATTCCTCCACCTCTCCCTTTCAGGAGAGAGCA 5597
 QY 1880 ArgProThrLysSerLysGlySerLysSerSerArgSerSerSerLeuGlyAsnLysSer 1899
 Db 5598 CGACCCACTAAAGCAAGGAGCAAAATCTAGTCGAAGCAGTTCTTGGGGAATTAAGC 5657
 QY 1900 ProGlnLeuSerGlyAsnLeuSerGlyGlnSerAlaAlaSerValLeuHisProGlnGln 1919
 Db 5658 CCCAGGCTTCAGGTAACTGTCTGTCTGAGAGTGCAGCTTCAGCTTCGACCCCGACAG 5717
 QY 1920 ThrLeuHisProProGlnValAsnIleProGlnSerGlyGlnAsnGlnLeuGlnProLeu 1939
 Db 5718 ACCCTCCACCTCTCGCAACATCCGAGTCGCGGAGATCGCTTACACGCCCTT 5777
 QY 1940 LysProSerProSerSerAspAsnLeuTyThrSerAlaPheThrSerAspGlyAlaIleSer 1959
 Db 5778 AAGCACTCTCCCTCCAGTACCACTCTATTCAGCTTCACAGTATGATGCGCATTTCA 5837
 QY 1960 ValProSerLeuSerAlaProGlyGlnGlyThrSerThrAsnThrValGlyAlaThr 1979
 Db 5838 GTCACCAAGCCTTCTCTGCTCCAGGTCAAGTAAATTAACCAACATCATCGCCAAAACAA 5897
 QY 1980 ValAsnSerGlnAlaAla 1985
 Db 5898 TAAATGAGAGATGTTGCC 5915
 PCT-US04-02652-453
 ; Sequence 453, Application PC/TUS0402652
 ; GENERAL INFORMATION:
 ; APPLICANT: CARGILL, Michele et al.
 ; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
 ; FILE REFERENCE: CLO01505
 ; CURRENT APPLICATION NUMBER: PCT/US04/02652
 ; CURRENT FILING DATE: 2004-01-30
 ; NUMBER OF SEQ ID NOS: 50231
 ; SOFTWARE: FASTSEQ for Windows Version 4.0
 ; SEQ ID NO 453
 ; LENGTH: 4872
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (1)...(4872)
 ; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1-
 PCT-US04-02652-453

Alignment Scores:

Pred. No.: 0 Length: 4872
 Score: 7154.00 Matches: 1425
 Percent Similarity: 97.67% Conservative: 0
 Best Local Similarity: 97.67% Mismatches: 6
 Query Match: 66.17% Indels: 28
 DB: 1 Gaps: 1

US-10-010-720-14 (1-2136) x PCT-US04-02652-453 (1-4872)

QY 678 GlySerGlnHisGlnGlnAlaHisSerThrGlyThrValProGlyHisIleProSerThr 697
 Db 3 GGTTCCCAACATGAAAGGACATTTCTACAGGACAGATCCAGGGGCAATAACCTTACT 62
 QY 698 ValGlnAlaGlnSerGlnProHisGlyValTyProProSerSerValAlaGlnGlyGln 717
 Db 63 GTCCAGCACAGTCTCAGCCCATGGGATATTCACCTTCAAGTGTG----- 110
 QY 718 SerGlnGlyInProSerSerSerSerLeuThrGlyValSerSerSerGlnProIleGln 737
 Db 110 ----- 110
 QY 738 HisProGlnGlnGlnGlnGlyIleGlnGlnThrAlaProProGlnGlnThrValGlnTy 757
 Db 111 -----CAGCAGGAAATACACACACAGCCCTCTCTCAACAGACAGTGCAGTAT 158
 QY 758 SerLeuSerGlnThrSerThrSerSerSerGlyAlaThrThrAlaGlnProValSerGlnPro 777
 Db 159 TCACCTTTCACAGACATCACTCCAGTACAGCCCACTCTCAGCAGCCAGTACGACCT 218
 QY 778 GlnAlaProGlnValLeuProGlnValSerAlaGlyLysGlnSerThrGlnGlyValSer 797
 Db 219 CAAGCTCCACAGATCTTGCCTCAAGTATCAGCTGGAAGAAACAGAGTACTCAGGAGTCTCT 278
 QY 798 GlnValAlaProAlaGlnProValAlaValAlaGlnProGlnAlaThrGlnProThrThr 817
 Db 279 CAGGTGCTCTCTCAGAGCCAGTTCAGTACAGTACAGCCCAAGCTACCGCCAGCCACT 338
 QY 818 LeuAlaSerSerValAspSerAlaHisSerAspValAlaSerGlyLysSerAspGlyAsn 837
 Db 339 TTGGCTTCTCTGTAAGACATGACATTCAGATGTTGCTTCAAGTATGATGATGCAAT 398
 QY 838 GlnAsnValProSerSerSerSerGlyArgHisGlnGlyValArgThrThrLysArgHisTyArg 857
 Db 399 GAGAACGTCCTCCATCTTCCAGTACAGGACGAGGAGAAAGAACTCAAAACGGCATTAACGA 458
 QY 858 LysSerValArgSerArgSerArgHisGlnLysThrSerArgProLysLeuAspGlyLeu 877
 Db 459 AAATCTGTAAGAGAGTCTCTGACATGAAAGAAATTCACGCCCAAAATTAAGAAATTTTG 518
 QY 878 AsnValSerAsnLysGlyAspArgValAlaGlyCysGlnLeuGlnThrHisAsnArgLys 897
 Db 519 AATGTTTCAATTAAGAGACCGAGTAGTAGATGCAATTAAGACTCATTAATAGGAAA 578
 QY 898 MetValThrPheLysPheAspLeuAspGlyAspAsnProGlnGlnIleAlaThrIleMet 917
 Db 579 ATGTTTACATTAATTAATTTGACCTAGATGTGACAAACCCGAGAGGATAGCAACATTAATG 638
 QY 918 ValAsnAsnAspPheIleLeuAlaIleGlnArgGlnSerPheValAspGlnValArgGln 937
 Db 639 GTGACACATGACTTATCTTACGAAATGAGAGAGAGCGTTTGATGATCAAGTCCGAGAA 698
 QY 938 IleIleGlnLysAlaAspGlnMetLeuSerGlnAspValSerValGlnProGlnGlyAsp 957
 Db 699 ATTATGAAAAAAGCTATGTAATGCTCAGTACAGAGATGCTAGTGAAACAGAGGGATGAT 758
 QY 958 GlnGlyLeuGlnSerLeuGlnGlyLysAspAspTyThrGlyPheSerGlySerGlnLysLeu 977
 Db 759 CAGGAGATGAGAGTCTACAAAGAGAGAGATGACTATGCTTTTCAAGGTTCTCGAAATTTG 818
 QY 978 GlnGlyGlnPheLysGlnProIleProAlaSerSerMetProGlnGlnIleGlyIlePro 997
 Db 819 GAAAGAGAGTTCAACCAACCAATTCCTGCTCTTCCATGCGCAACAGCAAAAGGATTCCT 878
 QY 998 ThrSerSerLeuThrGlnValAlaHisSerAlaGlyArgArgPheIleValSerProVal 1017
 Db 879 ACCAGTTCTTAACTCAAGTTGTTCAATTCGCGGAAAGGGGTTTATAGAGAGCTGCTG 928
 QY 1018 ProGlnSerArgLeuArgGlnSerLysValPheProSerGlnIleThrAspThrValAla 1037
 Db 939 CGAAGAAAGCGGATTAAGAGATCAAAAGTTTCCCGCAGTGAATAACAGATACAGATGCT 998

QY 1038 AlaSerThrAlaGlnSerProGlyMetAsnLeuSerHisSerAlaSerSerLeuSerLeu 1057
DB 999 GCCCTTAAGCTCAGAGCCCTGGAAATGAACTGTCTCACTCGATCATCTCCCTTAAGCTA 1058
QY 1058 GlnGlnAlaPheSerGluLeuArgAlaGlnMetThrGlnGlyProAsnThrAlaPro 1077
DB 1059 CAACAGGCTTTCTGAACTTAGAGCTGACGCTCCAAATGACAGAAAGAACCAANACAGCACT 1118
QY 1078 ProAsnPheSerHisThrGlyProThrPheProValValProPheLeuSerSerIle 1097
DB 1119 CCAACTTAGTACATACAGACCAACATTTCCAGTAAGTACTCTCTTTTAAGTAGACTT 1178
QY 1098 AlaGlyValProThrThrAlaAlaAlaThrAlaProValProAlaThrSerSerProPro 1117
DB 1179 GCTGAGGCTCCAAATACAGACGACCAACAGCAGCTCCCTGCAACAGAGCCCTCT 1238
QY 1118 AsnAspIleSerThrSerValIleGlnSerGluValThrValProThrGlnGluGlyIle 1137
DB 1239 AATGACATTTCCACATCTGATTAATTCAGTCTAGGTTACAGTGCCTGAAGAGGGAGATT 1298
QY 1138 AlaGlyValAlaThrSerThrGlyValValValThrSerGlyLeuProIleProProVal 1157
DB 1299 GCTGAGGTGCCACAGACAGAGGTGGTAATTCAAGTGGTCTCCCATACCACTGTG 1358
QY 1158 SerGluSerProValLeuSerSerValValSerSerIleThrIleProAlaValValSer 1177
DB 1359 TCTGAATCACCACTGACTTCCAGGTGATTCAAGTATCACAAATACCTGCAGTTGTCTCA 1418
QY 1178 IleSerThrThrSerProSerLeuGlnValProThrSerThrSerGluIleValValSer 1197
DB 1419 AATATCTACTACATCCCGTCACTTCAGATCCCAATCCACATCTGAGATCGTTGTTCT 1478
QY 1198 SerThrAlaLeuThrProSerValThrValSerAlaThrSerAlaSerAlaGlySer 1217
DB 1479 AGTACACACAGTATCTCTTCAAGTAAAGTTTCAAGCACTTCAGCTCTGACGGGGAGT 1538
QY 1218 ThrAlaThrProGlyProGlyProProAlaValValSerGlnGlnAlaGlySerThr 1237
DB 1539 ACTGCTACCCAGGATCTTAAGCTCCAGCTGATATCTCAGCGAGGACAGCACT 1598
QY 1238 ThrValGlyAlaThrLeuThrSerValSerThrThrThrSerPheProSerThrAlaSer 1257
DB 1599 ACTGTGGAGCCCATTTAACTCATGTTTCAACCACTTCATTCACCAAGCACTCA 1658
QY 1258 GlnLeuSerIleGlnLeuSerSerSerThrSerThrProThrLeuAlaGluThrValVal 1277
DB 1659 CAGCTGTCATTCAGCTTACAGCACTTCACTCTACTTACGTGAACCGTGTGTA 1718
QY 1278 ValSerAlaHisSerLeuAspLysThrSerHisSerSerThrThrGlyLeuAlaPheSer 1297
DB 1719 GTTAGCCCACTCATAGATTAAGACATCTCATAGCACTGATCACTGATGGCTTTCTCC 1778
QY 1298 LeuSerAlaProSerSerSerSerSerProGlyAlaGlyValSerSerThrIleSerGln 1317
DB 1779 CTCTCTGACCATTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1838
QY 1318 ProGlyGlyLeuHisProLeuValIleProSerValIleAlaSerThrProIleLeuPro 1337
DB 1839 CCTGTGGGTGATCTTGGTATTCATTCATGATAGTTTACTCTCATTTCTTCCC 1898
QY 1338 GlnAlaAlaGlyProThrThrThrProLeuLeuProGlnValProSerIleProProLeu 1357
DB 1899 CAAGCACACAGACCTACTTCTACACCTTTATTAACCAAGTACTAGTATCCACCTTG 1958
QY 1358 ValGlnProValAlaAsnValProAlaValGlnGlnThrLeuIleHisSerGlnProGln 1377
DB 1959 GTTACAGCTGTGCTGCAATGCTGCTGTATACAGACACACTAATTCATAGCAGCTCAA 2018
QY 1378 ProAlaLeuLeuProAsnGlnProHisThrHisCysProGluValAspSerAspThrGln 1397
DB 2019 CCAAGCTTTGGCTCCCAACAGCCCACTACTATGTCTGTGAAGTAAAGTTCTGATACCAA 2078
QY 1398 ProLysAlaProGlyIleAspAspIleLysThrLeuGlnGluLysLeuArgSerLeuPhe 1417

DB 2079 CCCAAAGCTCTGGAATGATGACATTAAGACTCTTAAGAAAGAGCTGCGTCTGTTC 2138
QY 1418 SerGlnHisSerSerSerGlyAlaGlnHisAlaSerValSerLeuGluThrSerLeuVal 1437
DB 2139 AGTAAACAGCTTATCTGAGGCTCAGACATGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2198
QY 1438 IleGluSerThrValThrProGlyIleProThrThrAlaValAlaProSerLysLeuLeu 1457
DB 2199 ATAGAGAGCACTGTACACACAGGATCCCACTACTCTGTGTGACCAAGAACTCCTG 2258
QY 1458 ThrSerThrThrSerThrCysLeuProProThrAsnLeuProLeuGlyThrValAlaLeu 1477
DB 2259 ACTTCTACCAAGTACTTCTTACTTACCAACCAATTTACCATAGGAACAGTGTCTTG 2318
QY 1478 ProValThrProValValThrProGlyGlnValValSerThrProValSerThrThrThrSer 1497
DB 2319 CCAATTACACAGTGTGTCACACTGGGGAAGTTTCAACCCAGTCAACACACTACTATCA 2378
QY 1498 GlyValLysProGlyThrAlaProSerLysProProLeuThrLysAlaProValLeuPro 1517
DB 2379 GAGTGAACCTGGAACCTGCTCCCTCAAGCACTCTTAAGCTCCGCTCTGCCA 2438
QY 1518 ValGlyThrGluLeuProAlaGlyThrLeuProSerGluGlnLeuProPheProGly 1537
DB 2439 GTGGGTACTGAACCTTCAGAGGATCTACCCAGCAGCAGCTGCCACCTTTTCCAGA 2498
QY 1538 ProSerLeuThrGlnSerGlnGlnProLeuGlnAspLeuAspAlaGlnLeuArgArgThr 1557
DB 2499 CTTTCTTAACCAAGTCCAGCAACCTTGAGAGATCTTAATGCTCAATTAAGAAAGACA 2558
QY 1558 LeuSerProGluLysIleThrValThrSerAlaValGlyProValSerMetAlaAlaPro 1577
DB 2559 CTTAGTCAAGATATKATACAGATGACTTGCAGGTGTCTGTCTGATGCGGCTCYA 2618
QY 1578 ThrAlaIleThrGluAlaGlyThrGlnProGlnLysGlyValSerGlnValLysGluGly 1597
DB 2619 ACAGCAATCAACAGAGGAGAACACAGCTCAGAGAGGTGTTTCTCAAGTCAAGAAAGAGC 2678
QY 1598 ProValLeuAlaThrSerSerSerGlyAlaGlyValPheLysMetGlyArgPheGlnValSer 1617
DB 2679 CCGTCTTACCACTACTTCAAGAGCTGGTGTTTAAAGTGGACCAATTCAGGTTTCT 2738
QY 1618 ValAlaAlaAspGlyAlaGlnLysGluGlyLysAsnLysSerGluAspAlaLysSerVal 1637
DB 2739 GTTGCACACACGGTGCAGAAAGAGGTAAATAAGTCAGAAAGTGCAAAGTCTGTT 2798
QY 1638 HisPheGluSerSerThrSerGluSerSerValLeuSerSerSerSerProGluSerThr 1657
DB 2799 CATTGTGAATCCAGACCTCAGAGTCTCTCAGTCTATCAAGTATGATCCAGAGATACC 2858
QY 1658 LeuValLysProGluProAsnGlyIleThrIleProGlyIleSerSerAspValProGlu 1677
DB 2859 TTGATGAACCAAGGCGGAATGGATTAACATCCCTGTGTATCTTCTTAAGTGTGCCAG 2918
QY 1678 SerAlaHisLysThrThrAlaSerGluAlaLysSerAspThrGlyGlnProThrLysVal 1697
DB 2919 AGTGCCCACAAACTACTGCTCAGAGGCAAAAGTCAAGCACTGGCAGCGCTACCAAGTT 2978
QY 1698 GlyArgPheGlnValThrThrThrAlaAsnLysValGlyArgPheSerValSerLysThr 1717
DB 2979 GAGAGTTTTCAGGTGACCACTACAGCAAAACAAAGTGGGTGTTTCTCTGTATCAAAAAC 3038
QY 1718 GluAspLysIleThrAspThrLysLysGluGlyProAlaAlaSerProPhePheAsp 1737
DB 3039 GAGACAAAGTACTGACAAAGAAAGAAAGAACAGCTGGCACTCTCTCTTTATGAT 3098
QY 1738 IleGlnGlnAlaValLeuProAlaValIleProLysLysGluLysProGluLeuSerGlu 1757
DB 3099 TTGGAACAAAGCTGTTCTCTCTGCTGTGATACAAAGAAABARAAGCTGAAGTGCAG 3158
QY 1758 ProSerHisLeuAsnGlyProSerSerAspProGluAlaAlaPheLeuSerArgAspVal 1777

Db 3159 CCTCACATCTAAATGGGCGCTCTTCGACCCGAGGCGCTTTTAAAGTAGGATGTG 3218
QY 1778 AspAspGlySerGlySerProHisSerProHisGlnLeuSerLysSerLeuProSer 1797
Db 3219 GATGATGATGTCGGGTAGTCCACACTCGCCCATCGAGCTCAAGAGCCTTCTTAGC 3278
QY 1798 GlnAsnLeuSerGlnSerLeuSerAsnSerPheAsnSerSerTyMetSerSerAspAsn 1817
Db 3279 CAGAACTTAAGTCAAAAGCCTTAGTAATTCATTAACTCTTACATGAGTAGCAACAT 3338
QY 1818 GluSerAspIleGluAspGluAspLeuLysLeuGlnLeuArgArgLeuArgAspLysHis 1837
Db 3339 GAGTCAGATATCGAAGATGAAGACTTAAAGTTAGAGTCGACGACTACGAGATAAAT 3398
QY 1838 LeuLysGluIleGlnAspLeuGlnSerArgGlnLysHisGluIleGluSerLeuTyThr 1857
Db 3399 CTCAAAGAGATTGAGACCTGACAGATGCGCAGAGACATGAATTTGATATACC 3458
QY 1858 LysLeuGlyLysValProProAlaValIleIleProProAlaAlaProLeuSerGlyArg 1877
Db 3459 AAACCTGGGCAAGGTGCCCCCTGCTGTATTATTCCTCCAGCTGCTCCCTTTACAGGAGA 3518
QY 1878 ArgArgArgProThrLysSerLysGlySerLysSerSerArgSerSerSerLeuGlyAsn 1897
Db 3519 AGACGACGACCCACTAAAGCAAGGAGGAGCAAACTTAGTCGAGCAGTTCTTGGGGAT 3578
QY 1898 LysSerProGlnLeuSerGlyAsnLeuSerGlyGlnSerAlaAlaSerValLeuHisPro 1917
Db 3579 AAAAGCCCCAGCTTTCAGGTACCTGCTGTCAGATGACGCTTCAGCTTCACCCCC 3638
QY 1918 GlnGlnThrLeuHisProProGlyAsnIleProGluSerGlyGlnAsnGlnLeuGln 1937
Db 3639 CAGCAGACCTCCACCTCTCGGCAACATCCCAAGTCCGGGCAAGATCAGCTGTACAG 3698
QY 1938 ProLeuLysProSerProSerSerSerAspAsnLeuTySerAlaPheThrSerAspGlyAla 1957
Db 3699 CCCCTTAAGCATCTCCCTCCAGTAGCAACTTATTACGCTTCACCAAGTAGTGTC 3758
QY 1958 IleSerValProSerLeuSerAlaProGlyGlnGlyThrSerSerThrAsnThrValGly 1977
Db 3759 ATTTCAGTACCAAGCTTTCGCTCCAGGTCAAGAACACAGCAGACACAAACATGTTGG 3818
QY 1978 AlaThrValAsnSerGlnAlaAlaGlnAlaGlnProProAlaMetThrSerSerArgLys 1997
Db 3819 GCACAGTGAACAGCCAGCCGCCCCAGCTCAGCTCTGCGCATGACCTCCAGCAGAGAG 3878
QY 1998 GlyThrPheThrAspAspLeuHisLysLeuValAspAsnTPalaArgAspAlaMetAsn 2017
Db 3879 GGCACATTCACAGATGACTTGCAACAAGTTGGTAGACAATGGCCCCGAGATGCCATGAT 3938
QY 2018 LeuSerGlyArgArgGlySerLysGlyHisMetAsnTyrgLysIleProGlyMetAlaArg 2037
Db 3939 CTCACAGCAGAGAGAGCAAGGCAAAAGGCAATGAGGGCCCTCGAATGGCAAG 3998
QY 2038 LysPheSerAlaProGlyGlnLeuCysIleSerMetThrSerAsnLeuGlyGlySerAla 2057
Db 3999 AAGTCTCTGACCTGGGCAACTGTGCACTCCATGACCTGGAACCTGGGTGGCTTGGC 4058
QY 2058 ProIleSerAlaAlaSerAlaThrSerLeuGlyHisPheThrLysSerMetCysProPro 2077
Db 4059 CCCATCTCTGAGATCAGTACTCTCTAGTCACTTCAACAAGTCTATGTCCCCCA 4118
QY 2078 GlnGlnTyrgLysPheProAlaThrProPheGlyAlaGlnTPSerGlyThrGlyGlyPro 2097
Db 4119 CAGCAGATAGGCTTTCAGTACCCCAATTTGGCGTCAATGAGTGGGACGGGTGGCCCA 4178
QY 2098 AlaProGlnProLeuGlyGlnPheGlnProValGlyThrAlaSerLeuGlnAsnPheAsn 2117
Db 4179 GCACCAAGCCACTTGGCCACTTCCAACTGTGGAACTGCTCTTGACAAATTTCAAC 4238
QY 2118 IleSerAsnLeuGlnLysSerIleSerAsnProProGlySerAsnLeuArgThrThr 2136
Db 4239 ATCAGCAATTTGCAGAATTCATCAGCAACCCCCAGGCTCCAACTGCGGACCACT 4295

Search completed: September 25, 2004, 13:36:14
Job time : 1955 secs